

**UNIVERSIDADE ESTADUAL DE CAMPINAS**  
**Faculdade de Engenharia Elétrica e Computação**  
**Departamento de comunicações**

**Marconi Pereira Fardin**

**Uma Alternativa à Inundação baseada em Tabelas de  
Dispersão para Planos de Controle de Redes Ópticas**

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Telecomunicações e Telemática.

Orientador:

Prof. Dr. Hélio Waldman – Unicamp

Co-orientador:

Prof. Dr. Moisés R. N. Ribeiro – UFES

Campinas, SP  
2008

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

F222a Fardin, Marconi Pereira  
Uma alternativa à inundação baseada em tabelas de dispersão para planos de controle de redes ópticas / Marconi Pereira Fardin. --Campinas, SP: [s.n.], 2008.

Orientadores: Helio Waldman, Moisés Renato Nunes Ribeiro.

Dissertação de Mestrado - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Redes de computação - protocolos. 2. Comunicações óticas. 3. Engenharia de tráfego. 4. Arquitetura de redes computador. 5. Algoritmos de computador. I. Waldman, Helio. II. Ribeiro, Moisés Renato Nunes. III. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. IV. Título.

Titulo em Inglês: A hash table based flooding alternative for optical networks control plane

Palavras-chave em Inglês: Protocols, computer network, Communications optical, Road traffic, Computer network architectures, Computer algorithms

Área de concentração: Telecomunicações e Telemática

Titulação: Mestre em Engenharia Elétrica

Banca examinadora: Carlos Alberto Kamienski, Maurício Ferreira Magalhães

Data da defesa: 04/11/2008

Programa de Pós Graduação: Engenharia Elétrica

**Marconi Pereira Fardin**

## **Uma Alternativa à Inundação baseada em Tabelas de Dispersão para Planos de Controle de Redes Ópticas**

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Telecomunicações e Telemática.

Prof. Dr. Hélio Waldman . . . . . Unicamp  
Prof. Dr. Carlos A. Kamienski . . . . UFABC  
Prof. Dr. Maurício F. Magalhães . . Unicamp


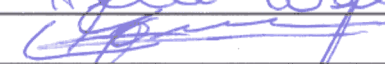
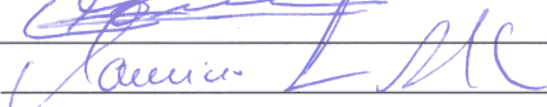
Campinas, SP  
2008

## COMISSÃO JULGADORA - TESE DE MESTRADO

**Candidato:** Marconi Pereira Fardin

**Data da Defesa:** 4 de novembro de 2008

**Título da Tese:** "Uma Alternativa à Inundação Baseada em Tabelas de Dispersão para Planos de Controle de Redes Ópticas"

Prof. Dr. Hélio Waldman (Presidente):   
Prof. Dr. Carlos Alberto Kamienski:   
Prof. Dr. Maurício Ferreira Magalhães: 

# Resumo

A infra-estrutura de transporte da internet está movendo-se para um modelo formado por roteadores de alta velocidade interconectados por um núcleo de redes ópticas inteligentes. A tecnologia WDM (*Wavelength Division Multiplex*) já está consolidada como meio de incrementar a capacidade de transmissão nessas redes porém, existem poucas propostas de planos de controles distribuídos capazes de lidar dinamicamente com as restrições de Qualidade de Serviço (QoS - *Quality of Service*) em grandes redes. Neste contexto, informações sobre os estados dos enlaces são constantemente divulgadas pela rede e o mecanismo de inundação se tornou ineficiente para lidar com a disseminação das informações. Este trabalho propõe e analisa uma estrutura sobreposta baseada em tabelas de dispersão distribuídas (DHT - *Distributed Hash Table*) para dispor informações em um espaço virtual. Uma rede sobreposta virtual endereçada por conteúdo (CAN - *Content Addressable Network*), com enlaces de longa distância para obter características de mundo pequeno, é utilizada para possibilitar o roteamento com baixa ocupação dos enlaces sem penalizar demasiadamente o atraso inserido. Avaliações analíticas ilustram o ganho de desempenho em grandes redes.

**Palavras-chave:** Plano de Controle Distribuído, Tabelas de Dispersão Distribuídas, Mundo Pequeno.

# Abstract

The transport layer structure of the Internet is moving towards a model of high-speed routers interconnected by intelligent optical core networks. A consensus has been reached on wavelength division multiplexing (WDM) as the way to increase the transmission capacity. However there are few proposals for distributed control planes under dynamic Quality of Service (QoS) constraints for large networks. In this context, resources availability must be frequently distributed and updated across the network; and flooding may no longer be the best strategy for information dissemination. This work proposes and assesses an overlay structure based on distributed hash tables (DHT) for information indexing over bi-dimensional virtual coordinated spaces. By adding long-distance interconnections to Content-Addressable Networks (CAN), the resulting *small-world* effect can reduce the number of hops per resource search. Analytical results indicate significant reduction in management traffic over the physical layer of large networks without penalizing to much the inserted delay.

**Keywords:** Distributed Control Plane, DHT, Small-World.

# Agradecimentos

À minha família e minha namorada pelo apoio durante o mestrado.

Ao meu orientador e co-orientador, Profs. Hélio Waldman e Moisés R.N. Ribeiro, sou grato pela oportunidade, ensinamentos e orientação.

À Faculdade de Engenharia Elétrica e Computação da pela estrutura disponibilizada.

Aos amigos Fernando Luis Trazzi Junior, Giuliano Salomão Sperandio e à todos da república 11HS pela força e pela descontração e oportunidade de convivência.

Aos colegas do Optinet, Wisstek e Labtel pelas críticas, sugestões e momentos de descontração.

Aos demais que de alguma forma contribuíram para a realização do mestrado

À FAPESP, pelo apoio financeiro.

*Aos meus pais*

# Sumário

<b>Lista de Figuras</b>	<b>xv</b>
<b>Lista de Tabelas</b>	<b>xvii</b>
<b>Glossário</b>	<b>xix</b>
<b>Trabalhos Publicados Pelo Autor</b>	<b>xxiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivações . . . . .	2
1.2 Trabalhos Relacionados . . . . .	3
1.3 Contribuições . . . . .	4
1.4 Organização do Texto . . . . .	5
<b>2 Redes Ópticas</b>	<b>7</b>
2.1 Histórico . . . . .	7
2.2 Redes Roteadas por Comprimento de Onda . . . . .	9
2.2.1 Componentes . . . . .	9
2.3 O Problema de GRWA . . . . .	11
2.4 Sumário . . . . .	14



---

<b>3</b>	<b>Plano de Controle, Sistemas P2P e o Efeito de Mundo Pequeno</b>	<b>17</b>
3.1	Plano de Controle de Redes Ópticas . . . . .	18
3.1.1	Protocolo de Descoberta de Recursos . . . . .	20
3.1.2	Protocolo de Gerenciamento de Informações . . . . .	22
3.1.3	Protocolo de Sinalização e Reserva . . . . .	24
3.2	Gerenciamento de Informações em Sistemas P2P . . . . .	25
3.2.1	Arquiteturas de Disseminação de Informações . . . . .	26
3.2.2	Sistemas Baseados em Tabelas de Dispersão Distribuídas . .	28
3.2.3	Funções de Dispersão . . . . .	29
3.2.4	Estrutura Semântica . . . . .	30
3.3	Redes de Mundo Pequeno . . . . .	36
3.3.1	Aplicações . . . . .	37
3.3.2	Reconfigurações Determinísticas . . . . .	39
3.4	Sumário . . . . .	40
<b>4</b>	<b>Mecanismo Proposto</b>	<b>43</b>
4.1	Plano de informação . . . . .	43
4.2	Exemplo de Construção do Plano de Informação . . . . .	48
4.3	Exemplo de Acesso às Informações . . . . .	53
4.4	Exemplo de Estabelecimento do Caminho Óptico . . . . .	55
4.5	Reconfiguração Determinística . . . . .	56
4.6	Sumário . . . . .	58
<b>5</b>	<b>Estudo Analítico</b>	<b>59</b>
5.1	Considerações Iniciais . . . . .	59
5.2	Atividades Analisadas . . . . .	60
5.3	Metodologia . . . . .	61

---

5.4	Equacionamento . . . . .	63
5.4.1	Atividade de Instalação . . . . .	63
5.4.2	Atividades de Operação . . . . .	65
5.4.3	Número Ótimo de Vizinhos . . . . .	67
5.5	Considerações Práticas . . . . .	68
5.6	Sumário . . . . .	68
<b>6</b>	<b>Análise dos Resultados</b>	<b>69</b>
6.1	Metodologia . . . . .	69
6.2	Topologia Física em Anel . . . . .	70
6.2.1	Carga por Enlace . . . . .	71
6.2.2	Número de Saltos . . . . .	74
6.3	Topologia física em Grade . . . . .	74
6.3.1	Carga por Enlace . . . . .	75
6.3.2	Número de saltos . . . . .	78
6.4	Sumário . . . . .	80
<b>7</b>	<b>Conclusão e Perspectivas</b>	<b>81</b>
	<b>Referências bibliográficas</b>	<b>84</b>

# Lista de Figuras

2.1	Técnicas de multiplexação. . . . .	8
2.2	Arquitetura de Rede Óptica. . . . .	9
3.1	Estrutura mantida pelo nó Pastry 10233102 para $b=2$ e $l=8$ . Todos os números estão na base 4 . . . . .	32
3.2	Tabela de apontadores no sistema Chord. . . . .	34
3.3	Estrutura CAN . . . . .	35
3.4	Formação de Rede de Mundo Pequeno através de reconfiguração aleatória. . . . .	38
3.5	Modelo de Kleinberg . . . . .	38
3.6	12 concentradores dispostos em espaço bi-dimensional para $d = 2$ e $n = 100$ . . . . .	39
4.1	Arquitetura dos nós. . . . .	44
4.2	Atividades de registro. . . . .	46
4.3	Atividades de estabelecimento de vizinhança. . . . .	47
4.4	Atividades para publicação de informação. . . . .	47
4.5	Atividades de busca de informação. . . . .	48
4.6	Redes adotadas. . . . .	49
4.7	Alterações na redes: Entrada do primeiro nó. . . . .	50

4.8	Alterações na redes: Entrada do segundo nó. . . . .	51
4.9	Alterações na redes: Entrada do terceiro nó. . . . .	52
4.10	Alterações na redes: Rede com todos os nós. . . . .	53
4.11	Rede do plano de informação formada pelas relações de vizinhanças através do TCP. . . . .	53
4.12	Reconfiguração determinística. . . . .	57
6.1	Topologia anel considerada na rede física . . . . .	70
6.2	Carga de instalação por enlace em função do tamanho da rede. . . .	71
6.3	Carga de instalação por enlace em função da quantidade de infor- mação a ser publicada. . . . .	72
6.4	Carga de Operação por enlace em função da quantidade de recursos a ser publicada. . . . .	73
6.5	Atraso para instalação do sistema. . . . .	74
6.6	Atraso na operação do sistema. . . . .	75
6.7	Topologia grade considerada na rede física . . . . .	76
6.8	Carga de instalação por enlace em função do tamanho da rede. . . .	76
6.9	Carga de instalação por enlace em função da quantidade de recursos a ser publicada. . . . .	77
6.10	Carga de Operação por enlace em função da quantidade de recurso a ser publicada. . . . .	78
6.11	Atraso para instalação do sistema. . . . .	79
6.12	Atraso na operação do sistema. . . . .	79

# Lista de Tabelas

4.1	Código dos Nós . . . . .	50
4.2	Código dos Recursos . . . . .	54

# Glossário

AS - *Autonomous System*

ASON - *Automatically Switched Optical Networks*

ATM - *Asynchronous Transfer Mode*

CAN - *Content Addressable Network*

CIA - *Centralized Indexing Architecture*

CSPF - *Constraint Shortest Path First*

DHT - *Distributed Hash Table*

DIFA - *Distributed Indexing Flooding Architecture*

DIHA - *Distributed Indexing Hashing Architecture*

FMP - *Fator de Mundo Pequeno*

Gbps - *Gigabits por segundo*

GMPLS - *Generalized Multi-Protocol Label Switching*

GRWA - *Grooming, Routing and Wavelength Assignment*

IETF - *Internet Engineering Task Force*

IP - *Internet Protocol*

IPTV - *TV over IP*

ITU-T - *International Telecommunication Union - Telecommunication Sector*

LMP - *Link Management Protocol*

LP - *Lightpath*

LSA - *Link State Advertisement*

LSDB - *Link State Data Base*

OLT - *Optical Line Terminals*

OSPF - *Open Shortest Path First*

OXC - *Optical Cross Connect*

P2P - *Peer-to-Peer*

QoS - *Quality of Service*

ROADM - *Reconfigurable Optical Add/Drop Multiplexer*

RSVP - *Reservation Protocol*

SDH - *Synchronous Digital Hierarchy*

SONET - *Synchronous Optical Network*

SW - *Small World*

TCP - *Transmission Control Protocol*

TDM - *Time Division Multiplexing*

TE - *Traffic Engineering*

VoIP - *Voice over IP*

WDM - *Wavelength Division Multiplexing*



# Trabalhos Publicados Pelo Autor

1. M. P. Fardin, M. R. N. Ribeiro, H. Waldman. “Small-World Reconfigured CAN for Distributed Control Plane of Future Optical Networks”. *XXVI Simpósio Brasileiro de Telecomunicações (SBrT’08)*, Rio de Janeiro, RJ, Brasil, Setembro 2008.

# Capítulo 1

## Introdução

O papel da informação no mundo atual é um dos mais fascinantes temas. Ela é o principal elemento para agregação de valor à produção de bens e serviços, e desempenha um papel estratégico e fundamental no processo de Inovação Tecnológica. Para tanto, a Informação é o canal que dá acesso ao conhecimento e que contribui para a mudança e o aperfeiçoamento, propiciando o conhecimento necessário à tomada de decisões e à execução de ações.

Um dos meios mais eficientes e utilizados para obter-se informação é a Rede Mundial de Computadores (Internet) e seu desenvolvimento é de grande importância nesta nova era. A Internet emergiu de uma rede militar da década de 60 que tinha como principal meta, em caso de destruição de um ponto da rede, estabelecer e manter a comunicação entre outros dois pontos quaisquer. Basicamente, isto se torna possível devido aos Protocolo de Internet (IP - *Internet protocol*) juntamente com outros protocolos baseados neste que possibilitam o encaminhamento confiável dos pacotes ao destino, através de vários roteadores (pontos da rede), por meio do endereço IP de forma distribuída. Este sistema não faz uso de nenhum mecanismo centralizado de gerencia. As informações sobre falhas são passadas entre os nós

e uma nova rota é estabelecida. Estes protocolos foram inicialmente especificados em [1, 2] sendo continuamente aperfeiçoados. Na década de 80 a Internet passou a interligar as universidades e somente na nos anos 90 chegou ao público geral, quando o volume de tráfego começou a apresentar um enorme crescimento, devido à popularização de serviços multimídia ultrapassando o tráfego de voz.

## 1.1 Motivações

O surgimento de novos serviços sobre a Internet como: voz sobre a rede IP (VoIP - *Voice over IP*), comércio eletrônico, compartilhamento de arquivos, televisão sobre rede IP (IPTV - *Television over IP*) juntamente com uma maior difusão da Internet nos domicílios faz com que crescimento do volume de tráfego de dados sobre esta rede cresça ainda mais rapidamente. Somado a isso, cresce também a exigência por Qualidade de Serviço (QoS - *Quality of Service*), das redes corporativas, para superar a tradicional política de menor esforço atualmente empregada. Contudo, a qualidade e a difusão desses serviços sobre a Internet têm sido limitadas pela capacidade de transmissão nos enlaces da rede e pela capacidade de gerência dos recursos presentes na rede.

Essas limitações têm despertado o interesse da comunidade científica que aprimoraram a capacidade de transmissão através da: substituição das redes metálicas por redes ópticas, aprimoramento de lasers semi-condutores, foto-detectores, grades sintonizáveis, matrizes de comutação óptica transparente e pelo desenvolvimento de avançadas técnicas de modulação que permitem a transmissão em uma única fibra centenas de canais e mais de 100 GB/s em cada um [3, 4, 5, 6].

Com advento destes nos novos recursos, a infraestrutura da Internet segue agora a tendência de um modelo formado por roteadores de alta velocidade conectados

por um núcleo de redes ópticas [7]. Neste contexto as redes ópticas transparentes com multiplexação por divisão de comprimento de onda (WDM - *Wavelength Division Multiplexing*) são as candidatas mais promissoras para o núcleo das redes de comunicação. Essa tecnologia tem permitido aumentar a capacidade de tráfego sem investimentos significativos na infra-estrutura instalada, o que vem ao encontro do crescente tráfego de dados na internet bem como à crescente necessidade de qualidade de serviço requerida pelas novas aplicações.

Para gerenciar esta nova estrutura e tentar garantir QoS novas funcionalidades foram incorporadas ao OSPF - *Open Shortest Path First* [8] em sua extensão para engenharia de tráfego (TE - *Traffic Engineering*) [9], possibilitando a inserção de restrições na escolha dos caminhos (CSPF - *Constraint Shortest Path First*) [10]. Todavia, ao manterem o mecanismo de inundação do OSPF, não se preocuparam com a quantidade de tráfego gerada pela necessidade constante de atualização das informações de controle para atender aos requisitos de qualidade de serviço em redes dinâmicas que interligam à camada eletrônica a camada óptica.

## 1.2 Trabalhos Relacionados

Outros trabalhos propõem técnicas para reduzir a quantidade de tráfego de controle e a mais antiga técnica utilizada é a hierarquização da rede dividindo a rede em áreas autônomas (AS - *Autonomous Systems*) [11], porém, torna o roteamento entre áreas sub-ótimo devido à restrição de informações de recursos de outras áreas. É proposto também, aumentar o intervalo entre as atualizações [12, 13], o que visivelmente aumenta a possibilidade de tornar a base de dados dos nós desatualizada podendo fazer com que os pacotes fiquem presos em rotas fechadas (*loops*). Estas propostas reduzem o acesso a informação tornando o sistema menos eficiente na

resolução do problema de agregação de tráfego, roteamento e alocação de comprimento de onda (GRWA - *Grooming, Routing and Wavelength Assignment*).

Recentemente, as técnicas de distribuição de informação utilizadas em aplicações par-a-par (P2P - *Peer-to-Peer*) foram utilizadas em planos de controle de redes IP. O uso de estratégias das aplicações P2P para complementar os protocolos GMPLS - *Generalized Multiprotocol Label Switching* aprimorando o provisionamento de rotas entre domínios pelos usuários [14]. Este trabalho considera apenas a comunicação entre domínios que se dá através de um sistema de gerência chamado de DOM (*Distributed Optical Manager*) o qual é baseado em tabelas de dispersão distribuídas (DHT - *Distributed Hash Table*). Em outro trabalho [15], é proposta a utilização de uma rede estruturada, também baseada em DHT, para oferecer serviços de sinalização.

Estes trabalhos apenas tratam da localização dos nós e não de informações para roteamento e baseiam na topologia Chord [16] em que a tabela de roteamento dos nós na rede overlay é definida e seu tamanho cresce com o aumento do número de nós, afetando a escalabilidade do sistema.

## 1.3 Contribuições

Neste trabalho propomos e avaliamos pela primeira vez a utilização de tabelas de dispersão distribuídas na rede de controle endereçada por conteúdo (CAN - *Content Addressable Network*) [17]. O uso deste mecanismo, consagrado nas aplicações P2P para disponibilizar informações atualizadas e unívocas aos nós da rede sem a utilização do mecanismo de inundação, tem como objetivo reduzir a quantidade de tráfego de gerência na rede em um cenário no qual existem muitos nós e os recursos são atualizados frequentemente. Para garantir simultaneamente a redução

do diâmetro da rede CAN e portanto, redução do tráfego e um menor tempo de busca na rede sobreposta propomos também reconfiguração determinística nas conexões desta rede sobreposta inspiradas no fenômeno de Mundo Pequeno (SW - *Small-World*). Um estudo analítico do atraso, e da banda utilizada na rede é realizado. Isto permite a comparação entre os mecanismos propostos e o baseado em inundação.

## 1.4 Organização do Texto

No Capítulo 2 abordamos as redes ópticas. Um breve histórico dos sistemas ópticos é apresentado; introduzimos as redes roteadas por comprimento de onda e seus principais componentes e, por fim, destacamos os principais desafios na resolução dos problemas de agregação de tráfego, roteamento e alocação de comprimento de onda. No Capítulo 3 abordamos o controle de grandes quantidades de informação. Iniciamos com o plano de controle de redes ópticas descrevendo brevemente os principais protocolos definidos pelo GMPLS. Apresentamos os sistemas desenvolvidos para aplicações P2P além de introduzirmos o fenômeno Mundo Pequeno. Neste contexto, um método de reconfiguração determinística de redes para obtenção deste fenômeno é descrito para obter uma maior escalabilidade no sistema. No Capítulo 4 propomos um mecanismo distribuído a ser utilizado pelo plano de controle para disponibilizar informação. Introduzimos com detalhes o funcionamento do sistema e o uso de reconfiguração determinística para obtenção de características de Mundo Pequeno para melhorar o seu desempenho. Um estudo analítico é apresentado no Capítulo 5 para a carga nos enlaces e o atraso inserido para modelo proposto e para o modelo clássico baseado em inundação. No Capítulo 6 analisamos os resultados numéricos para topologia em anel e em grade, comparando-os com o modelo de inundação presente nos sistemas atuais. No Capítulo 7 concluímos o trabalho e

apresentamos as perspectivas futuras.

## Capítulo 2

# Redes Ópticas

As redes ópticas provêem uma grande capacidade de transporte de tráfego nas redes que possibilita o oferecimento de uma grande variedade de serviços. Por estes motivos, estas redes tornaram-se o meio de transmissão preferido para cobrir longas distância e oferecer serviços em altas taxas. Neste capítulo apresentaremos um breve histórico das redes ópticas, descreveremos alguns de seus principais componentes e introduziremos o problema de agrupamento de tráfego, roteamento e atribuição de comprimento de onda (GRWA - *Grooming, Routing and Wavelength Assignment*).

### 2.1 Histórico

Dependendo do autor, há várias formas de classificação para as redes ópticas. Neste trabalho utilizaremos a definição apresentada por Ramaswami [18] que divide em duas gerações a evolução de tais redes.

A primeira geração teve início nos anos 70 e as redes ópticas eram usadas em enlaces ponto-a-ponto e sua principal atribuição era aumentar a capacidade de



transmissão do sistema. Durante esta geração, as funções de chaveamento, bem como toda inteligência da rede, estavam atribuídas ao plano eletrônico da rede, sendo necessária a conversão do sinal óptico em sinal eletrônico para a execução destas funções.

A técnica de multiplexação por divisão do tempo (TDM - *Time Division Multiplexing*), ilustrada na Figura 2.1(a), consiste em combinar sinais de taxas menores em um único sinal de taxa maior. Posteriormente, o desenvolvimento da multiplexação por divisão de comprimento de onda (WDM), ilustrada na Figura 2.1(b), que consiste em transmitir em uma única fibra vários comprimentos de ondas diferentes; isso possibilitou o aumento da taxa de transmissão por enlace. Neste cenário, o esquema de processamento eletrônico tornou-se incapaz de tratar todo o tráfego do nó, fazendo necessária a implantação de serviços de comutação com inteligência no próprio plano óptico da rede. Assim, cada nó tende a processar eletronicamente apenas o tráfego destinado a si próprio. Fato possibilitado pela capacidade de estabelecimento de conexões transparentes (i.e. sem processamento eletrônico intermediário) entre os nós. Tal estrutura é conhecida como segunda geração ou redes roteadas por comprimento de onda.

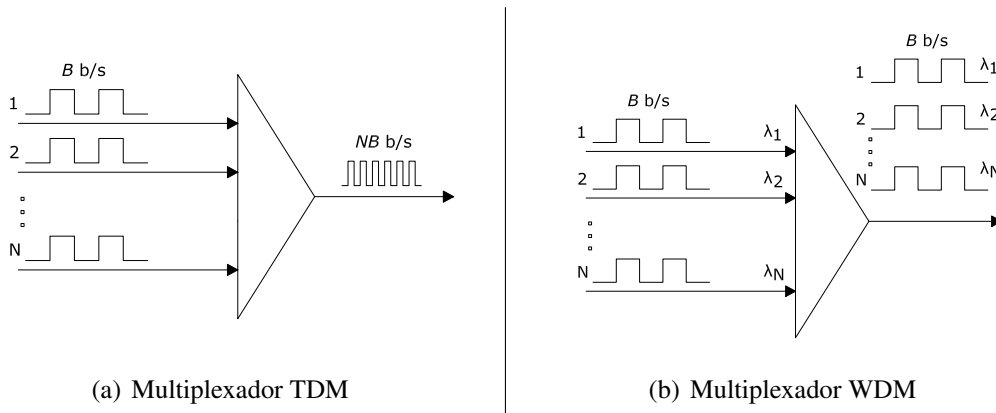


Figura 2.1: Técnicas de multiplexação.

Figura 2.2: Arquitetura de Rede Óptica.

## Fibra Ópticas

Os primeiros experimentos com fibra óptica foram realizados nos anos 60 nos quais conseguiram demonstrar que informações codificadas em sinais ópticos poderiam ser transmitidas através de um guia de onda de vidro fabricado através de processos especiais para garantir baixa atenuação. De forma bem simplista, uma fibra óptica é um cilindro de vidro muito fino composto por duas camadas denominadas núcleo e casca. Estas camadas são projetadas para guiar o sinal óptico entre dois dispositivos através do efeito da reflexão total.

A baixa degradação imposta ao sinal óptico pelas fibras tornou-as um dos meios de transmissão mais eficientes quando comparado com o espaço livre, nos sistemas sem fio, e os cabos de cobre. Em geral, as fibras têm como parâmetro de fabricação o seu coeficiente de atenuação  $\alpha$ , que chega até 0,2 dB/km. Além disso, os baixos níveis de dispersão permitem que altas taxas de transmissão sejam alcançadas em longas distâncias

O uso combinado de TDM e WDM permite que vários sinais distintos trafeguem em um mesmo comprimento de onda e que vários comprimentos de onda sejam transmitidos em uma única fibra.

## OLTs, ROADMs e OXCs

Um OLT multiplexa múltiplos comprimentos de onda em uma fibra e demultiplexa um conjunto de comprimentos de onda em uma única fibra em diversos sinais e podem retirar alguns desses sinais para uma aplicação local enquanto outros sinais podem apenas serem encaminhados a outros enlaces. Os OLTs são utilizados na extremidade de um enlace WDM ponto-a-ponto.

Os ROADMs lidam com sinais de múltiplos comprimentos de onda e podem, assim como os OLTs, retirar alguns comprimentos de onda para aplicações locais,

como também podem adicionar novos comprimentos de onda aos sinais ópticos. Os ROADMs também podem incorporar a capacidade de conversão de comprimento de onda.

Os OXCs, basicamente, realizam as mesmas funções dos ROADMs, porém em maior escala podendo comutar comprimentos de uma porta para outra, possibilitando a formação de redes com topologia em malha enquanto ROADM possibilitam apenas a topologia em anel.

Para lidar com demandas dinâmicas de tráfego, esses componentes têm desenvolvido a capacidade de rápida reconfiguração. Para que tal tarefa seja desempenhada de forma eficiente, é necessário que informações sobre as capacidades disponíveis nos diversos elementos da rede sejam de conhecimento de todos os nós da rede para que esses sejam bem utilizados no atendimento dinâmico das demandas.

## **2.3 O Problema de GRWA**

A escolha da rota a ser utilizada em um LP e os comprimentos de ondas atribuídos a cada enlace é resolvida através de algoritmos de roteamento e atribuição de comprimento de onda. Os algoritmos são comumente divididos na literatura entre estáticos e dinâmicos, sendo estes últimos os apropriados para atender a uma matriz de tráfego desconhecida ou com modificações muito freqüentes.

O objetivo de um problema de roteamento e atribuição de comprimento de onda é aprimorar o desempenho da rede em relação a probabilidade de bloqueio e a utilização da rede. Os algoritmos são divididos em duas categorias: as de tráfego estático e as de tráfego dinâmico. Como o tráfego predominante nas redes ópticas se tornou o IP, que não se consegue prever ou modelar, um algoritmo dinâmico se

faz necessário.

Geralmente a capacidade de transmissão do LP é muito maior que a necessidade de um único fluxo de tráfego. Por isso dedicar um LP para cada solicitação individual é ineficiente. A agregação de tráfego é inserida neste contexto para permitir que múltiplos fluxos de tráfego compartilhem um mesmo LP reduzindo o desperdício de capacidade disponível dos comprimentos de onda.

Além das restrições de disponibilidade de recursos a serem observadas na escolha da rota e na atribuição do comprimento de onda, um LP pode não ser estabelecido se restrições de camada física não forem atendidas. Efeitos intrínsecos do meio de transmissão que distorcem o sinal propagado devem ser considerados [19]. Além de efeitos clássicos como filtragem imperfeita dos canais WDM e saturação de ganho de amplificadores, costuma-se fazer uma divisão didática dos demais fenômenos físicos em duas categorias: lineares e não-lineares.

Os efeitos lineares, descritos abaixo, são independentes da potência do sinal e não causam interferência em outros canais.

**ASE - *Amplified Spontaneous Emission*:** Essas emissões são amplificadas nos dispositivos subsequentes junto com o sinal óptico que transporta informação. Embora os fótons emitidos estejam na mesma faixa de energia que os fótons incidentes do sinal óptico, eles são emitidos em direção, polarização e fase aleatórios. Assim, em redes ópticas de grandes distâncias, que se valem do cascadeamento de amplificadores ópticos para compensar as perdas do sistema, têm o acúmulo da ASE como o fator mais crítico a ser gerenciado.

**Dispersão Cromática:** Acontece devido ao fato da variação do índice de refração de acordo com a frequência. LPs de longas distâncias podem ser penalizados em virtude dos vários enlaces que são transpassados.

**PMD - Polarization-Mode Dispersion:** Acontece porque as fibras não são perfeitamente simétricas, de modo que dois modos ortogonalmente polarizados tenham diferentes constantes de propagação, causando um alargamento temporal do pulso. Em particular, sistemas em altas taxas de transmissão (10 ou 40 Gb) são afetados por este efeito.

Já os efeitos não-lineares são fortemente relacionados aos níveis de potência e causam interferência entre os canais WDM.

**SRS - Stimulated Raman Scattering:** No espalhamento estimulado de Raman o efeito tem alcance de até 15 THz. Neste caso, a onda que funciona como bombeio é um sinal de menor comprimento de onda e a onda de Stokes é um sinal de maior comprimento de onda. O canal de bombeio tem sua potência depletada enquanto o sinal com maior comprimento de onda apresenta ganho.

**SPM - Self-Phase Modulation:** No caso da auto-modulação de fase os sinais transmitidos sofrem gorjeio (chirping). Isto acontece porque o índice de refração da fibra é, na realidade, dependente da intensidade do sinal óptico. Esse índice de refração não-linear causa um deslocamento de fase que é proporcional à intensidade do pulso. Como diferentes partes do pulso possuem diferentes deslocamentos de fase, surge o gorjeio dos pulsos. O gorjeio induzido pela SPM pode aumentar bastante a largura do pulso devido à dispersão cromática que transforma a modulação de fase em variações de amplitude.

**XPM - Cross-Phase Modulation:** Nos sistemas WDM, os efeitos não-lineares são aumentados uma vez que os sinais combinados de todos os canais podem ter uma potência muito alta. Dessa forma, como o deslocamento de fase depende da intensidade do pulso com o conseqüente chirping induzidos pelo SPM, no XPM tem-se esse efeito causado pela intensidade dos outros canais.

**FWM - Four-Wave Mixing:** Em um sistema WDM que consiste no conjunto de frequências  $f_1, \dots, f_n$ , a mistura de quatro ondas gera novos sinais nas frequências com  $2f_i - f_j$  e  $f_i + f_j - f_k$ . Essa interferência entre canais é particularmente severa quando o espaçamento entre os canais é pequeno em fibras com baixa dispersão cromática.

Uma grande variedade de algoritmos para resolução do problema de GRWA são propostos na literatura, [20, 21, 22, 23, 24, 25], porém o desempenho alcançado por estes reduz à medida que impomos restrições às informações da rede. Isto torna necessária a publicação de diversas informações:

- a respeito de todos os comprimentos de onda e banda remanescente disponíveis em cada fibra da rede;
- comprimentos de onda passíveis de serem convertidos, combinações de chaveamentos possíveis entre as portas de cada OLT, OXC e OADM da rede;
- restrições de camada física para cada enlace.

Assim, para que cada nó possa atender as requisições dinâmicas, o conhecimento do estado destes recursos é essencial para o GRWA.

## 2.4 Sumário

O atendimento das demandas de forma dinâmica requer a constante atualização de todos os recursos disponíveis na rede para que informação desatualizada não ocasione em uma resolução do algoritmo de GRWA de baixo desempenho ou inconsistente.

Assim, para que cada nó possa atender as requisições dinâmicas, o conhecimento do estado destes recursos é essencial para resolução do problema de GRWA.

Portanto torna-se imprescindível o desenvolvimento de um plano de controle com um eficiente mecanismo de distribuição de informação.



## **Capítulo 3**

# **Plano de Controle, Sistemas P2P e o Efeito de Mundo Pequeno**

O plano de controle é responsável por gerenciar todos os recursos em uma rede óptica. Ele que torna disponível as informações da rede, torna possível a criação de conexões fim a fim e o gerenciamento de QoS destas bem como a proteção e restauração em caso de falhas. Neste Capítulo iremos abordar o plano de controle das redes ópticas e as Aplicações P2P, destacando o mecanismo de distribuição de informação em cada um deles. Apresentamos ainda o processo de reconfiguração determinística para redução do numero médio de saltos (diâmetro) entre os nós da rede sobreposta. O objetivo deste Capítulo é demonstrar similaridades entre o problema de distribuição de informação dos planos de controle e o sistema de disseminação/busca de informações de sistemas P2P. Esta contextualização é essencial para fundamentar o mecanismo proposto apresentado no Capítulo 4.

### 3.1 Plano de Controle de Redes Ópticas

A crescente demanda por gerenciamento nas redes IP direcionou o surgimento de novas funcionalidades nos seus roteadores. As operações nestes se tornavam cada vez mais complexas quando comparados com switches de camada 2, em particular Ethernet, e apresentavam um desempenho (avaliado em função da capacidade de encaminhamento) cada vez menor. Classicamente, ATM (*Asynchronous Transfer Mode*) era utilizada para transportar tráfego IP em altas velocidades e longas distâncias e as redes das operadoras tipicamente possuíam uma arquitetura composta por quatro redes distintas sobrepostas:

- IP para atender aplicações e serviços;
- ATM para controle de banda, monitoramento de desempenho e garantia de QoS (*Quality of Service*) na interconexão de dispositivos IP;
- SONET/SDH (*Synchronous Optical Network/Synchronous Digital Hierarchy*) para oferecer um transporte de alta qualidade para a carga paga sobre o meio físico, monitoramento de erros, OAM (*Operations, Administration and Management*) e proteção;
- WDM (*Wavelength Division Multiplexing*) para capacidade bruta de transmissão.

Esta arquitetura multicamadas tornou problemático lidar com a configuração, gerência e peculiaridades de cada camada, as interfaces de acesso entre camadas e a manutenção dos equipamentos distribuídos entre os diversos níveis. Além disso, a eficiência do transporte é comprometida devido ao *overhead* de sobrecarga de operação de cada protocolo. Nestas redes, o provisionamento de novos serviços

é uma tarefa complexa que requer um prazo excessivamente longo. Os recursos para o novo circuito devem ser reservados manualmente em cada camada (*hop-by-hop*). Após o estabelecimento do circuito, ainda é necessário uma fase de testes para verificar se a conexão está de acordo com os parâmetros descritos em SLAs (*Service Level Agreement*) [26].

Provedores de serviços, atualmente, tentam diminuir o número de camadas de suas redes sem comprometer as funcionalidades do modelo anterior buscando um novo modelo de operação. Acredita-se que o tráfego IP será transportado diretamente sobre redes ópticas roteadas por comprimento de onda [27]. Esta estrutura de rede é conhecida como IP sobre WDM [28]. A interligação entre o plano de gerenciamento e a rede óptica deve ser feita por um plano de controle que incorpore funções de chaveamento e de inteligência. Entre as principais funções destes planos de controle encontram-se:

- **Descoberta de Recursos:** inclui a descoberta automática de enlaces, vizinhança, correlações das propriedades dos enlaces e manutenção dos canais de sinalização;
- **Gerenciamento de Informação:** engloba as funções de divulgação dos dados descobertos pela função anterior e da topologia dentro do domínio de controle.
- **Sinalização e Reserva:** inclui as funções de provisionamento e restauração de circuitos.

Deste modo é possível o atendimento de demandas dinâmicas de tráfego nessas redes através do estabelecimento automático de circuitos ópticos.

O ITU-T (*International Telecommunication Union-Telecommunication Standardization Sector*) e a IETF (*Internet Engineering Task Force*) estão realizando um

esforço coordenado para padronizar um plano de controle inteligente e interfaces de acesso para as redes de transporte. Foi estabelecida pelo ITU-T uma arquitetura para o plano de controle de redes de transporte ópticas inteligentes chamada ASON (*Automatically Switched Optical Network*) [29], que define os requisitos e mecanismos de controle e gerenciamento em um alto nível de abstração, independente de tecnologia e de protocolo. Um conjunto de protocolos para exercer as funções do plano de controle chamado GMPLS (*Generalized Multi-Protocol Label Switching*) [30] também está em desenvolvimento pela IETF.

### 3.1.1 Protocolo de Descoberta de Recursos

Nas redes ópticas os elementos de rede adjacentes podem estar conectados por uma grande quantidade de enlaces. Utilizando-se da tecnologia WDM cada enlace por sua vez pode conter dezenas e até centenas de comprimentos de onda. A configuração manual das informações de identificação e capacidades de cada canal de dados presente em cada elemento da rede é inviável, por ser uma tarefa demorada, altamente sujeita a erros e inflexível em relação a mudanças. Paralelo a isto, existe uma separação entre os planos de controle e de transporte em redes GMPLS, e estes planos não necessariamente possuem a mesma topologia ou compartilham o mesmo meio físico. Devido a estes motivos, o IETF introduziu o protocolo LMP (*Link Management Protocol*) [31], responsável pela manutenção dos canais de controle, descoberta de vizinhança e correlação das propriedades dos enlaces. O protocolo LMP é um protocolo ponto-a-ponto que utiliza o protocolo UDP (*User Datagram Protocol*), ou seja, mensagens LMP são trocadas apenas por elementos de rede adjacentes. Para estabelecer uma relação de adjacência LMP, pelo menos um canal de controle deve estar ativo entre um par de nós GMPLS. O LMP é encarregado de realizar uma série de funcionalidades distintas, entre elas [32]:

- Durante o processo de inicialização, vizinhos adjacentes trocam mensagem LMP para ativar o canal de controle e definir suas identidades e capacidades. Uma vez que o canal de controle é ativado, sua manutenção é realizada através de trocas regulares de mensagens *Hello*;
- O processo de Descoberta de Enlaces (*Link Discovery*) auxilia um LSR a determinar a existência, a conectividade e a natureza de seus enlaces. Inicialmente, a única informação disponível sobre os enlaces são seus identificadores locais. Seus estados de operação são ignorados e os identificadores remotos (identificadores usados pelos LSRs nas outras extremidades dos enlaces) são desconhecidos. As mensagens trocadas durante o processo de descoberta de enlaces resolvem estas questões;
- O processo de Troca de Capacidade de Enlaces (*Link Capabilities Exchange*) é uma etapa opcional, em que LSRs trocam informações sobre características específicas dos enlaces de dados, como por exemplo parâmetros de Engenharia de Tráfego (TE - *Traffic Engineering*);
- A Verificação dos Enlaces (*Link Verification*) pode ser conduzida a qualquer instante para verificar os estados dos enlaces de dados entre dois pares LMP. Os processos de verificação são idênticos aos processos de descoberta dos enlaces;
- O Isolamento de Falhas (*Fault Isolation*) é uma das mais importantes características do LMP, particularmente em redes fotônicas. Nestas redes, os circuitos ópticos são compostos por vários enlaces e, na eventual falha de um enlace (por exemplo devido a um corte de fibra), todos os nós da rota no sentido *downstream* detectarão a perda do sinal. Por isso, os métodos tradicionais de monitoramento do estado operacional de enlaces (*Link Health*

*Monitoring*) de camadas 2 e 3 não são mais apropriados. O LMP introduz um sistema baseado em mensagens capaz de isolar de forma eficiente o ponto exato de falha de um circuito.

### 3.1.2 Protocolo de Gerenciamento de Informações

O protocolo de gerenciamento de informação mais utilizado atualmente tem sido o OSPF (*Open Shortest Path First*). Seu desenvolvimento teve início em 1987 motivado pelas falhas apresentadas pelo seu antecessor o protocolo RIP (*Routing Information Protocol*). O RIP é baseado em vetor de distância, onde a métrica usada para calcular as rotas é a distância das sub-redes. Por sua vez, OSPF é um protocolo baseado em estado de enlace, que utiliza uma métrica mais flexível. Um custo é atribuído a cada enlace, usualmente associado com a largura de banda nominal. Além de eliminar os problemas do RIP, o OSPF introduziu novas funcionalidades como multi-caminho de custo equivalente (*equal-cost multipath*), hierarquia de roteamento, separação de rotas internas e externas e uma segurança aperfeiçoada.

Um roteador OSPF divulga as informações do estado de seus enlaces através de LSAs (*Link State Advertisements*). LSAs são distribuídos na rede através da inundação que garante que todos os roteadores em uma rede (ou área OSPF) terão a mesma base de dados, chamada de LSDB (*Link State DataBase*). Para possibilitar a manutenção da consistência das LSDB cada LSA carrega informações de controle em um cabeçalho de 20 bytes, e são necessárias atualizações constantes.

Os LSAs são carregados entre roteadores encapsulados em pacotes IP. Os endereços de destino são sempre os identificadores IP dos roteadores OSPF que estabeleceram previamente uma relação de vizinhança com o roteador de origem, ou o endereço multicast caso o domínio de roteamento seja uma rede *broadcast*.

O roteamento em cada roteador OSPF é realizado aplicando um algoritmo SPF

(*Shortest Path First*) em uma árvore construída a partir da LSDB. Este processo resulta em caminhos para cada sub-rede conhecida, e o próximo salto para cada destino alcançável é inserido na tabela de roteamento. Porém, todo tráfego destinado para uma mesma sub-rede usará sempre a mesma rota, mesmo que não haja largura de banda suficiente nos enlaces que compõem a rota para transportar todo tráfego, causando problemas de congestionamento. O grupo de trabalho do IETF propôs aperfeiçoamentos no protocolo OSPF para permitir a divulgação de métricas que possibilitem a realização de Engenharia de Tráfego (TE - *Traffic Engineering*) [9].

A inclusão de métricas adicionais a serem divulgadas pelos LSAs como banda disponível e atraso do enlace, são propostas em [10] e algoritmos de roteamento com restrições (CSPF - *Constrained Shortest Path First*) são previstos para viabilizar o provisionamento de QoS. Todavia, a frequência da divulgação de alterações das métricas surge como um possível limitante prático; principalmente no contexto do uso de inundação para LSA. A proposta de extensão do OSPF para inclusão de engenharia de tráfego mantém inalterada a forma de LSA via inundação [9].

Lembrando que ainda há grupo grande de métricas específicas das redes ópticas a serem incluídas [26], a frequência de atualização e volume de tráfego de LSA circulando no plano de controle pode ser um entrave severo à escalabilidade, podendo exigir restrição ao número de nós devido a um problema bem conhecido que diz que em uma rede de  $N$  nós são necessários  $N^2$  LSAs para atualizar a base de dados de todos os nós [11].

Muitas propostas presentes na literatura atual concordam que um dos principais entraves à escalabilidade do plano de controle de redes ópticas seria o mecanismo de inundação utilizado para propagar informação a respeito dos estados dos enlaces, pois ele consome muita capacidade da rede [12, 13, 33, 34, 35, 36].

Estas propostas investigam diversas maneiras de mitigar os problemas apresentados por este mecanismo, porém não resolvem o problema da escalabilidade das redes, apenas tornam possível um crescimento um pouco maior propondo uma redução da informação disponibilizada aos nós, reduzindo a probabilidade de se escolher um caminho ótimo para um melhor funcionamento da rede [37].

### 3.1.3 Protocolo de Sinalização e Reserva

O protocolo de sinalização e reserva mais utilizado em redes GMPLS é o RSVP-TE. Este protocolo é responsável basicamente pelo processo de troca de mensagens no plano de controle para estabelecer, manter, modificar e remover LPs. Extensões ao RSVP são propostas para suportar engenharia de tráfego conhecido como RSVP-TE [38]. As principais características do RSVP-TE são:

- Uso de mensagens PATH e RESV para solicitar e atribuir rótulos durante o processo de criação de LSPs;
- A capacidade de especificar uma rota explícita durante o estabelecimento ou re-roteamento de LSPs;
- A capacidade de especificar largura de banda e outros parâmetros de QoS durante o estabelecimento de LSPs;
- Um novo protocolo de *Hello* para manter relações de adjacência entre pares RSVP-TE.

A eficiência e confiabilidade de operações de engenharia de tráfego, como a agregação dinâmica de demandas atendendo a requisitos de QoS, decrescem à medida que são impostas barreiras ao conhecimento do estado atual da rede [39]; seja por conhecimento restrito da topologia e recursos disponíveis ou por informações



desatualizadas [40]. Identificamos, assim, o mecanismo de disseminação de informações no plano de controle como um ponto chave para o projeto do plano de controle de redes GMPLS mas que não vem recebendo a devida atenção. Um indício de tal problema é o efeito de “*route flapping*” observado em redes IP sob OSPF. Ele é resultado da atualização do estado do enlace (LSU - *link state update*) atingir os roteadores da área em tempos diversos, levando a cálculos dos caminhos com menor custo acumulado, e conseqüentes atualizações de tabelas de roteamento, a serem feitos de forma assíncrona. Até o momento onde todos os roteadores atualizam suas tabelas (convergência), os pacotes - incluindo os de atualização do estado do enlace - podem ficar presos em rotas fechadas (“*loops*”) ou seguir fora da rota de menor custo [41].

Porém, as aplicações P2P, responsáveis pela maioria do tráfego na Internet, atravessaram o mesmo problema de escalabilidade e também de falta de segurança em seu mecanismo de distribuição de informação sendo solucionado com a proposta de uma nova arquitetura que possui um mecanismo estruturado de alocação e busca de informação. Sistemas baseados em DHT se apresentam bastante evoluídos e são propostos para toda sorte de aplicações distribuídas presentes na internet. Estes sistemas motivam um estudo de sua aplicação como mecanismo de distribuição de informação de planos de controle óptico que, como já mostrado, vem enfrentando problemas com o atual mecanismo de distribuição de informação baseado em inundação.

## 3.2 Gerenciamento de Informações em Sistemas P2P

Recentemente, semelhanças entre os problemas enfrentados pelos protocolos de gerenciamento de informação do plano de controle de redes ópticas e de sistemas

P2P foram identificadas. Assim, as soluções de escalabilidade adotadas em sistemas P2P estão sendo propostas para incorporação ao plano de controle de redes ópticas. Uma estrutura hierárquica sobreposta é proposta em [15] para substituir protocolos de sinalização em planos de controle de redes ópticas. Em outra proposta, o uso de estratégias desenvolvidas em sistemas P2P são aplicadas na resolução de aprovisionamento de recursos em roteamento entre domínios feitos por clientes [14].

Os sistemas P2P de compartilhamento de arquivos tornaram-se populares com as aplicações como Napster e Gnutella devido à grande capacidade de distribuir informações presentes nessas redes, atraindo assim o interesse científico para o desenvolvimento de novas aplicações baseadas nesse princípio.

Podemos dividir a evolução dessas aplicações em três arquiteturas: as de índice centralizados (CIA - *Centralized Indexing Architecture*), as de índice distribuído com inundação (DIFA - *Distributed Indexing with Flooding Architecture*) e as de índice distribuído com dispersão (DIHA - *Distributed Indexing with Hashing Architecture*) conforme descrevemos abaixo.

### 3.2.1 Arquiteturas de Disseminação de Informações

- **CIA** - Uma aplicação P2P com uma arquitetura CIA é aquela que contém um servidor central ou um cluster de servidores que é responsável por responder os pedidos de busca e realizar todas as tarefas de manutenção da infraestrutura.

O exemplo mais comum de uma aplicação deste tipo é o sistema Napster [42]. Existe nestas arquiteturas um ponto central que se estiver inoperante, a rede não funciona. Redes deste tipo não são verdadeiramente P2P, porque apenas a transferência de arquivos é feita de forma distribuída. O mecanismo de buscas e a manutenção da infra-estrutura são realizados de forma centralizada,

conforme o paradigma cliente/servidor.

- **DIFA** - Esta arquitetura é caracterizada pela completa descentralização de seu funcionamento. Os mecanismos de busca e manutenção da infra-estrutura estão distribuídos pela rede. Neste tipo de sistema, cada nó é responsável por manter a listagem dos seus próprios arquivos e responder quando receber uma busca. Para isto, no DIFA, é utilizado o mecanismo de inundação. A aplicação deste tipo mais conhecida e estudada é a rede Gnutella [43].

As aplicações deste tipo possuem uma limitação muito grande que está ligada ao fato de seu mecanismo de busca utilizar o mecanismo de inundação. Para que uma rede não fique saturada com a inundação de mensagens, estas possuem um número máximo de nós que podem ser atravessados. Isto possui uma séria implicação: mesmo que um arquivo exista no sistema e que o nó que o contenha esteja disponível, uma busca para este arquivo pode falhar.

Sistema semelhante é utilizado nos planos de controle de redes ópticas, onde a divulgação da informação também é feita utilizando a inundação. O problema da saturação também é enfrentado pelos protocolos de gerenciamento de informação. A divisão da rede em sistemas autônomos (ASs - *Autonomous Systems*) com o objetivo de impedir que informações de redes proprietárias sejam de domínio público limitam a inundação impedindo o acesso de alguns nós às informações do sistema podendo resultar em rotas sub-ótimas.

- **DIHA** - Esta arquitetura também possui uma característica totalmente descentralizada. A principal diferença entre as redes DIFA e DIHA está no mecanismo de busca. Nos sistemas com inundação, cada par é responsável pelo espaço de índices relativo aos arquivos que ele próprio contém. A arquitetura DIHA muda este conceito, cada nó é responsável por um subconjunto do

espaço total de índices. Quando o nó entra na rede, este recebe um espaço do conjunto dos índices dos arquivos, ao sair da rede o mesmo deverá alocar estes índices em outro nó. As buscas não são difundidas na rede sem direção como na DIFA, ao invés disto, são direcionadas para o nó correto que é o responsável pelo respectivo índice dentro do espaço de índices.

Uma grande quantidade de aplicações utilizarem esta arquitetura como Bit-Torren [44], eMule [45], e diversos softwares de mensagens instantâneas [46, 47]; ela começou com projetos de pesquisa como Plaxton [48], Chord [16], Tapestry [49], Pastry [50] e CAN [17]. Estes projetos são baseados em Tabelas de Dispersão Distribuída (DHT -*Distributed Hash Table*) [51] e são diferenciados pela geometria de roteamento usada [52].

Esses sistemas são dinâmicos e configurados automaticamente, têm busca direcionada (escalável) e armazenamento de informações distribuído (sobrevivência a falhas). Características estas, desejáveis aos planos de controle de grandes redes uma vez que são conseguidas sem restrição de informação a elemento algum da rede.

### 3.2.2 Sistemas Baseados em Tabelas de Dispersão Distribuídas

Tabelas de dispersão distribuídas provêm os mesmo serviços de busca de tabelas de dispersão (*hash table*). Através de uma função de distribuição (*hash function*) mapeia-se os recursos em códigos e armazena-se as tuplas <Código, ID> que será buscado através do código gerado. A diferença entre os mecanismos é que no primeiro as tuplas são armazenadas em diversos pontos da rede.

Este mecanismo foi desenvolvido inicialmente para aplicações P2P de compartilhamento de arquivos mas já foi adaptado para aplicações P2P de mensagens

instantâneas, voz sobre IP (VOIP) entre outras e também tem sido proposta para redes ad-hoc móveis [53] e é utilizado no mapeamento e na busca de recursos em um espaço semântico distribuído entre os nós da rede.

Sistemas DHT utilizam uma única função básica que é a localização de um código ( $h$ ) e retornar o identificador (ID) relacionado a este - Busca( $h$ ). As principais características destes sistemas são: a descentralização, em que os nós agem coletivamente sem a necessidade de nenhum elemento centralizador; a escalabilidade, que permite um sistema eficiente com centenas de milhares de nós; e a tolerância a falhas, que permite o funcionamento confiável do sistema mesmo com entradas, saídas e falhas de nós na rede.

### 3.2.3 Funções de Dispersão

Tabelas de dispersão distribuídas são uma estrutura de dados que associam os recursos (*chaves*) compartilhados em códigos (*valores*) de pesquisas. Seu objetivo é, a partir de uma chave, em um procedimento rápido, encontrar o valor associado. O mapeamento das chaves ( $k$ ) em valores ( $h$ ) é feito por uma função de dispersão ( $H$ ).

$$H(k) = h \quad (3.1)$$

O ideal para as funções de dispersão é que sejam sempre fornecidos índices únicos para as chaves de entrada. A função perfeita seria a que, para quaisquer entradas A e B, sendo A diferente de B, fornecesse saídas diferentes. Quando as entradas A e B são diferentes e, passando pela função de dispersão, geram a mesma saída, acontece o que chamamos de colisão.

Na prática, funções de dispersão perfeitas ou quase perfeitas são encontradas

apenas onde a colisão é intolerável – criptografia – ou quando conhecemos previamente o conteúdo da tabela armazenada. Nas tabelas de dispersão comuns a colisão é apenas indesejável, diminuindo o desempenho do sistema. Muitos programas funcionam sem que seu responsável suspeite que a função de dispersão seja ruim e esteja atrapalhando o desempenho. Há diversos algoritmos de resolução de colisão [54], mas os mais conhecidos são Encadeamento Separado e Endereçamento Aberto.

Apesar das tabelas de dispersão terem em média tempo constante de busca, o tempo gasto no desenvolvimento é significativo. Avaliar uma boa função de dispersão é extremamente custoso e profundamente relacionado à estatística. Na maioria dos casos soluções mais simples como uma lista encadeada devem ser levadas em consideração. Existem várias funções de código aberto disponíveis entre eles SHA - *Secure Hash Standard* [55], MD5 - *Message-Digest Algorithm* [56] e WHIRLPOOL especificado na ISO 10118-3.

É importante ressaltar que em uma aplicação distribuída de tabela de dispersão, todos os pares da rede precisam possuir a mesma função de dispersão.

### 3.2.4 Estrutura Semântica

Uma função de dispersão mapeia recursos compartilhados em códigos. Porém, para um funcionamento distribuído de uma tabela de dispersão é necessária uma estrutura para possibilitar a distribuição deste mapa entre os nós e a comunicação entre eles. O espaço semântico é onde cada elemento da rede se registra e publica seus recursos a serem compartilhados. É importante ressaltar que as posições dos recursos e dos nós no espaço semântico correspondem a um mapeamento semântico, podendo não ter relação alguma com a posição física dos nós.

O espaço semântico em uma tabela de dispersão distribuída é dividido entre os

nós da rede assim que estes entram na rede. Um par, ao tornar-se disponível na rede gera um código, através de uma função de dispersão, de si próprio e então gera uma requisição de entrada na rede. Esta requisição é roteada através do espaço semântico a partir de um ponto de conhecimento prévio, até o ponto definido pelo código. Ao chegar neste ponto o nó negocia com o responsável pelo ponto a divisão do espaço e as relações de vizinhanças a serem estabelecidas.

Após esta operação, o nó divide o espaço semântico e fica encarregado de gerenciar o novo sub-espço criado. Nestes sub-espços criados serão publicados os recursos a serem compartilhados.

A publicação de um recurso consiste também em gerar um código, através da função de dispersão, para o recurso a ser inserido, roteá-lo, com a função *busca(h)*, até o ponto de inserção, que estará em um sub-espço de algum par, e inserir na tabela de dispersão armazenada neste nó a tupla  $\langle \text{Código}, ID \rangle$  contendo o código gerado (*Código*) e o identificador (*ID*) do nó que detém o recurso, responsável pela publicação.

A comunicação entre os pares se dá através de uma rede sobreposta formada, basicamente, através das relações de vizinhanças entre os pares. Cada par, ao se registrar na rede, torna-se responsável por um subespço e faz ligações TCP com os vizinhos imediatos no espaço semântico (nós mais próximos).

O mecanismo de roteamento é a chave para o bom funcionamento desses sistemas. Nessas redes, o roteamento acontece como uma linha reta da origem até o destino através da rede sobreposta. Este mecanismo apenas repassa o pacote, sem multiplicá-lo em várias direções, de vizinho em vizinho até que se atinja o destino.

Um nó ao gerar uma requisição verifica se o destino está em seu subespço, caso contrário, ele verifica qual de seus vizinhos está mais próximo ao destino final e passa a requisição. Esta operação é repetida até que a requisição atinja seu destino

final.

Um nó nessa rede mantém uma pequena tabela lógica de roteamento que mantém o endereço IP e a coordenada do subespaço virtual de seus vizinhos imediatos no espaço coordenado. As aplicações desenvolvidas se diferenciam justamente pela maneira de definir seus vizinhos e serão brevemente descritas abaixo.

- **Pastry**

Cada nó neste sistema é relacionado a um código de 128 bits. Este código é utilizado para definir a posição em um espaço semântico circular em que as posições variam de 0 a  $2^{128} - 1$ . Cada nó mantém um tabela de roteamento (*Routing table*), um conjunto de vizinhos (*Neighborhood Set*) e um conjunto de folhas (*Leaf Set*) como observado na Figura 3.1.

NodeID 10233102			
Conjunto de Folhas		MENOR	MAIOR
10233033	10233021	10233120	10233122
10233001	10233000	10233230	10233132

Tabela de Roteamento			
-0-2212102	1	-2-2301203	-3-1203203
0	1-1-301233	1-2-230203	1-3-021022
10-0-31203	10-1-32102	2	10-3-23302
102-0-0230	102-1-1302	102-2-2302	3
1023-0-322	1023-1-000	1023-2-121	3
10233-0-01	1	10233-2-32	
0		102331-2-0	
		2	

Conjunto de Vizinhos			
13021022	10200230	11301233	31301233
02212102	22301203	31203203	33213321

Figura 3.1: Estrutura mantida pelo nó Pastry 10233102 para  $b=2$  e  $l=8$ . Todos os números estão na base 4

A tabela de roteamento de um nó é organizada em  $\lceil \log_{2^b} n \rceil$  linhas com  $2^b - 1$  entrada em cada. As  $2^b - 1$  entradas na linha  $n$  referem-se a nós que tem o



código com os mesmos bits nos primeiros  $N$  dígitos e tem um dos  $2^b - 1$  possíveis valores para o  $[N + 1]$ -ésimo dígito. O identificador (ID) do nó também é armazenado. É escolhido o prefixo numericamente mais próximo do nó de destino.

A escolha do valor  $b$  obedece a uma relação de eficiência entre armazenamento em um nó e número de saltos entre nós, quanto maior for  $b$  maior será a tabela de roteamento e conseqüentemente menor o número médio de saltos entre origem e destino.

O conjunto de folhas mantém  $L$  códigos que são os  $L/2$  maiores e os  $L/2$  menores numericamente mais próximos. Valores típicos de  $L$  são  $2^b$  ou  $2 \times 2^b$ .

Para realizar o roteamento primeiramente é verificado se o código de destino está contido no conjunto de folhas, caso positivo o pacote a ser roteado é repassado diretamente ao nó de destino. Caso contrário o pacote é repassado para o nó com identificador numericamente mais próximo do destino.

- **Chord**

Cada nó neste sistema armazena dois conjuntos de vizinhos. A lista de sucessores que contém os sucessores mais próximos do nó  $i$  e a lista de apontadores (*finger list*) que são os sucessores dos nós  $[(i + 2^{k-1}) \bmod 2^m]$  para o nó  $i$  uma rede com código de  $m$  bits como ilustrado na Figura 3.2.

O roteamento consiste em passar o pacote a ser roteado para o nó com o código numericamente mais próximo do destino porém não maior. A precisão do roteamento é garantida pela lista de sucessores e a eficiência pela lista de apontadores.

Este sistema é o mais utilizado em trabalhos que propõem o uso das técnicas P2P. Os trabalhos que propõem o uso das técnicas P2P em plano de controle

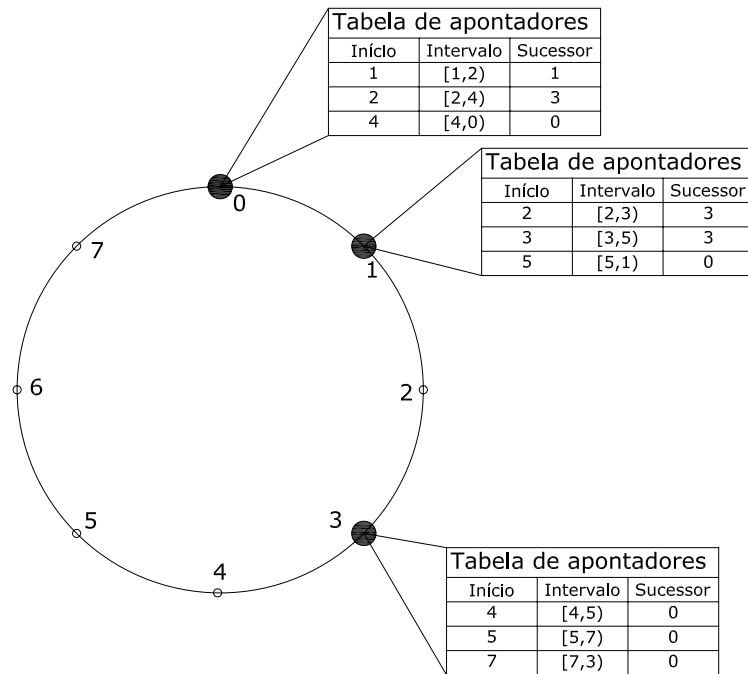


Figura 3.2: Tabela de apontadores no sistema Chord.

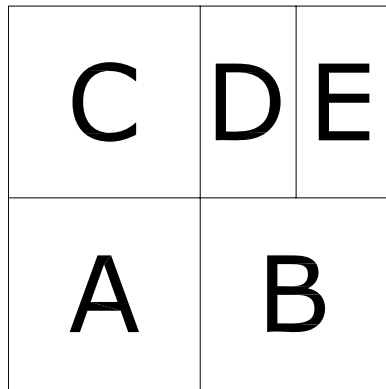
[15, 14], também fazem uso deste sistema.

- **CAN**

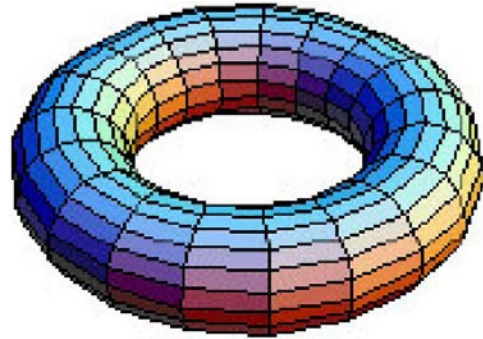
Este sistema é baseado em uma espaço virtual  $d$ -dimensional de coordenadas cartesianas sobre um toróide. Este espaço é completamente lógico e não possui nenhuma relação com as coordenadas física dos nós. O espaço coordenado é completamente repartido entre os nós de forma que cada nó possua seu próprio subespaço virtual.

Na figura 3.3 a estrutura CAN é mostrada. Na Figura 3.3(a) um espaço bi-dimensional dividido entre 5 nós no qual o nó  $C$  tem como vizinhos os nós  $A$ ,  $D$  e o nó  $E$  que também compartilha aresta quando em um toróide. Por simplicidade não mostramos o toróide nesta figura, porém, a união das extremidades do espaço mostrado formam uma superfície toroidal mostrado

na Figura 3.3(b).



(a) Espaço bi-dimensional com 5 nós



(b) Toróide

Figura 3.3: Estrutura CAN

Este espaço coordenado virtual é utilizado para armazenar pares (Chave, Valor). Para armazenar o nó  $(K_1, V_1)$  a chave  $K_1$  é mapeada, de forma determinística, em um ponto  $P$  do espaço coordenado utilizando uma função de dispersão. O nó  $(K_1, V_1)$  é armazenado no nó responsável pelo subespaço ao qual pertence o ponto  $P$ . Se o ponto  $P$  não pertence ao nó que deseja armazenar o par, a requisição necessita ser roteada através da estrutura CAN até atingir o nó responsável pelo subespaço ao qual pertence o ponto  $P$ . Portanto, o roteamento é uma tarefa crítica no CAN.

Os nós CAN são organizados automaticamente na rede sobreposta que representa o espaço semântico virtual. Um nó descobre e mantém o endereço IP dos nós com os quais suas arestas são compartilhadas. O endereço deste conjunto de vizinhos imediatos no espaço virtual possibilita o roteamento entre dois pontos quaisquer.

O roteamento acontece de maneira intuitiva seguindo diretamente da origem ao destino. Este consiste em repassar a requisição ao vizinho no espaço virtual

mais próximo do destino especificado pelo ponto  $P$ .

Em um espaço  $d$ -dimensional repartido igualmente em  $n$  subespaços idênticos o número médio de saltos entre dois nós é dado por  $(d/4)(n^{1/d})$  cada nó mantém uma tabela de  $2d$  vizinhos. Isto mostra que, em um espaço virtual  $d$ -dimensional, podemos aumentar o número de nós sem aumentar a quantidade de informação armazenada em cada nó resultando em um aumento do número de saltos na ordem de  $(n^{1/d})$ .

Algumas propostas tentam melhorar a eficiência do mecanismo de roteamento em estruturas P2P [57, 58] utilizando o fenômeno de Mundo Pequeno (SW - *Small-World*), oriundos de estudos de psicologia social. Investigaremos neste trabalho a aplicabilidade deste efeito em estruturas com dimensões maiores que 1 como em CAN.

### 3.3 Redes de Mundo Pequeno

Uma rede social exibe características de Mundo Pequeno se qualquer dois indivíduos nessa rede estiverem ligados por um pequeno número de pessoas conhecidas. Os estudos sobre esse fenômeno começaram na década de 60 conduzidos por Stanley Milgram [59]. Trabalhos recentes atestam que o fenômeno é verificado em várias redes, tanto na natureza quanto em redes desenvolvidas pelo homem [60, 61, 62]. O experimento de Stanley Milgram seguiu uma maneira intuitiva de se pensar sobre o fenômeno. O objetivo era verificar que um pequeno número de pessoas interligam duas pessoas quaisquer, dentro do território dos Estados Unidos da América. Em seus testes uma pessoa (fonte) foi escolhida em Nebraska para mandar uma carta para uma pessoa (destino) em Massachusetts. A fonte sabia apenas o nome e a profissão do destino e tinha que passar a carta para uma pessoa

conhecida pelo seu primeiro nome, de modo que a carta chegasse mais rapidamente ao destino. A pessoa que recebesse a carta receberia as mesmas informações e repassaria a carta até que esta chegasse ao destino. Foi observado que as cartas passavam, em média, por 5 a 6 pessoas, ficando conhecido como o princípio dos 6 graus de separação.

### 3.3.1 Aplicações

As Redes de Mundo Pequeno foram introduzidas por Watts e Strogatz como modelo de redes de distribuição elétrica, difusão de doenças na população, mercado de capitais e muitos outros inclusive a Internet [60]. Nesse estudo eles demonstram a origem do fenômeno através de uma rede em anel, onde cada nó deste anel estaria ligado aos nós vizinhos a uma distância menor que uma pré-determinada. A partir disso se percorre todo o anel e verificando se haverá reconfiguração das conexões através de uma probabilidade também pré-determinada ( $p$ ). As redes de Mundo Pequeno seriam um meio termo entre a aleatoriedade total e uma rede regular como pode ser visto na Figura 3.4. É mostrado em [60] que para  $p = 0,20$  temos uma redução de 80% no diâmetro (i.e. número médio de saltos entre nós) da rede e para  $p$  maiores o diâmetro pouco muda. Este fenômeno se caracteriza pelo fato de que a reconfiguração de poucas ligações reduzirem drasticamente o diâmetro da rede.

John Kleinberg propôs um modelo onde os nós não estariam em anel e sim em uma grade  $n$ -dimensional [63] (semelhante à CAN) e estariam ligados aos nós com uma distância pré-determinada,  $a$ , como pode ser visto na Figura 3.5(a) com  $a = 1$ . Além destas conexões de curta distância, alguns nós podem ter ligações de longas distancias aleatórias, como pode ser visto na Figura 3.5(b) onde umas das conexões do nó  $u$  é com o nó  $v$ .

Todavia, estes trabalhos apresentam métodos empíricos de reconfiguração, sendo

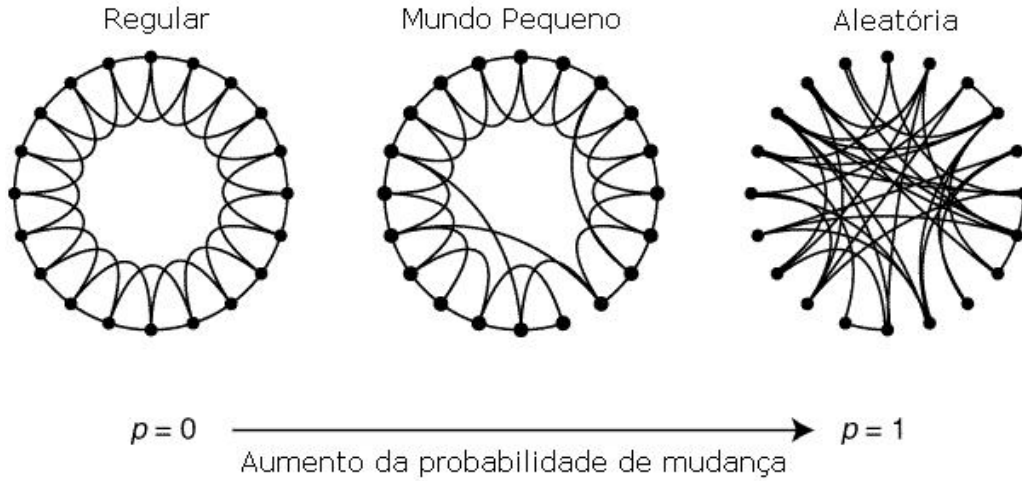


Figura 3.4: Formação de Rede de Mundo Pequeno através de reconfiguração aleatória.

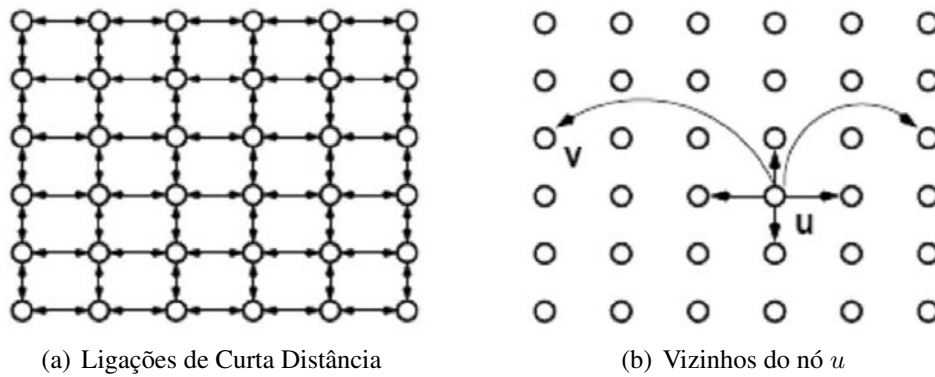


Figura 3.5: Modelo de Kleinberg

as ligações de longas distâncias conhecidas mediante sorteio, não garantindo o diâmetro da configuração final.

### 3.3.2 Reconfigurações Determinísticas

Francesc Comellas propôs uma reconfiguração determinística do anel apresentado anteriormente na Figura 3.4 [64]. A proposta consiste na eleição de alguns nós como concentradores, e apenas estes fazem ligações de longa distância. Há ainda a criação de uma nova rede que conecta estes concentradores. A grande vantagem deste modelo é saber quantas e quais ligações devem ser feitas para se reduzir o número de saltos no caminho crítico a um valor determinado.

Fardin estende este modelo unidimensional para redes  $n$ -dimensionais [65]. Este modelo elege certo número de nós da rede para tornarem-se concentradores, isto é, fazer ligações de longa distância, de acordo com o diâmetro desejado,  $D$ . Para uma dada rede de interligação com diâmetro  $D_H$ , podemos obter  $d$ . Na Figura 3.6 temos essa grandeza ilustrada em um espaço bi-dimensional.

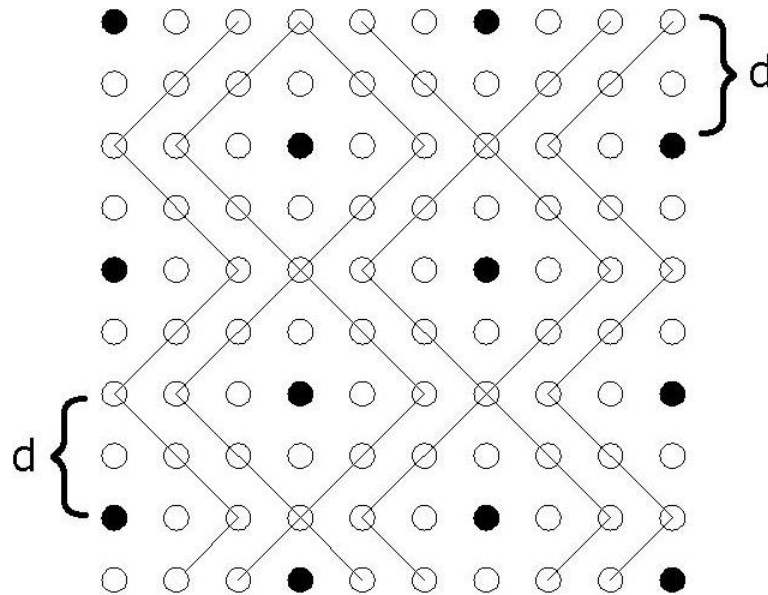


Figura 3.6: 12 concentradores dispostos em espaço bi-dimensional para  $d = 2$  e  $n = 100$ .

$$d = \left\lceil \frac{D - D_H}{2} \right\rceil \quad (3.2)$$

Sendo  $N$  o número de nós em uma coluna ou linha do sistema, o número de concentradores,  $N_H$ , pode ser obtido através da equação 3.3.

$$N_H = \left\lceil \frac{N - 1}{2d + 2} \right\rceil \left\lceil \frac{N - 1}{2d} \right\rceil + \left\lceil \frac{N - d - 2}{2d + 1} + 1 \right\rceil \left\lceil \frac{N - d - 1}{2d + 1} \right\rceil \quad (3.3)$$

A reconfiguração determinística de redes CAN possibilita a redução do diâmetro. Este mecanismo mostra-se uma opção interessante na redução do tráfego e do atraso na implantação de um sistema de gerenciamento de informações de redes multi-camadas baseadas em sistemas P2P.

### 3.4 Sumário

O plano de controle de redes ópticas se apresenta uma arquitetura composta por quatro camadas. Este empilhamento multi-camada não é satisfatório, tornando a gerência problemática. Um novo modelo conhecido como IP sobre WDM foi proposto e é o principal candidato para gerência dessas redes. Um plano de controle que inclua atividades de gerenciamento de enlaces e informação juntamente com mecanismos de reservas de recursos foi desenvolvido através de um esforço coordenado do ITU-T e IETF. Porém, o uso de inundação para divulgar informações do estado da rede torna a proposta inviável.

Sistema P2P também atravessaram problemas de escalabilidade devido ao uso do mecanismo de inundação. Nestes sistemas, o problema foi resolvido através do uso de DHT que permite a disponibilização estruturada de informações. O fenômeno de Mundo Pequeno é utilizado em conjunto com o uso de DHT para



aumentar a escalabilidade desses sistemas. Reconfigurações determinísticas de topologia para obtenção de características de Mundo Pequeno são propostas para facilitar a configuração e análise do sistema.

## Capítulo 4

### Mecanismo Proposto

Neste capítulo propomos um sistema de gerenciamento de informação que consiste no desenvolvimento de um plano de informação sobreposto ao plano de controle composto por estrutura semântica semelhante à CAN. Serão apresentados também neste capítulo as operações que permitem o funcionamento do sistema e um exemplo para elucidar o funcionamento. Por fim apresentaremos um mecanismo de reconfiguração determinística de topologia para o sistema.

#### 4.1 Plano de informação

As alterações propostas para o plano de controle óptico para torná-lo mais eficiente e confiável para as operações de engenharia de tráfego e para agregação dinâmica de demandas atendendo a requisitos de QoS consistem em evoluí-lo de uma arquitetura baseada em inundação para uma DIHA.

Para fazermos esta evolução necessitamos mudar o paradigma de distribuição da informação, deixando de distribuí-la arbitrariamente, via inundação, para busca-la quando necessária. Na Figura 4.1 expomos a arquitetura atual do plano de controle

4.1(a) e apresentamos as mudanças propostas em 4.1(b). Nota-se, que a diferença entre os dois modelos consiste no acesso às informações da rede. Enquanto no modelo atual, o protocolo de roteamento com restrições (CSPF - *Constraint Shortest Path First*) busca informações em uma base de dados por ele armazenada, o modelo proposto, através de uma função de dispersão, irá buscar as informações no plano de informação.

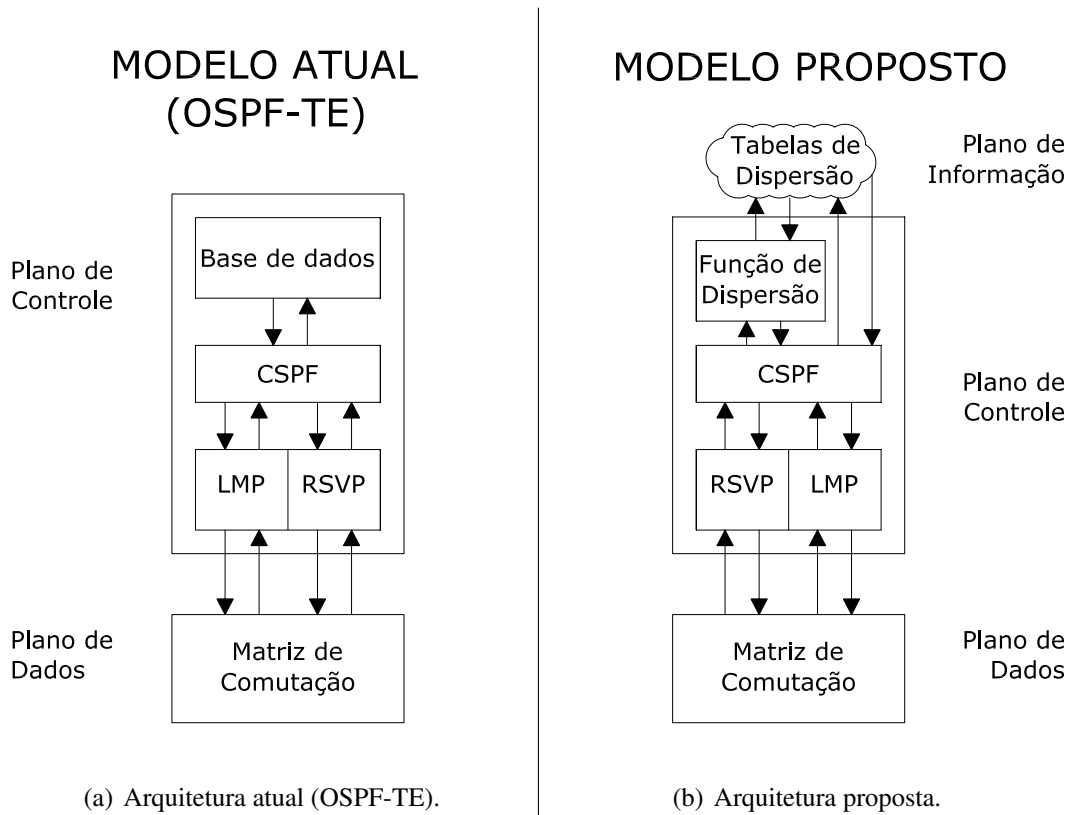


Figura 4.1: Arquitetura dos nós.

O plano de informação é composto por nós interligados por uma rede de informação. Os nós do plano de informação serão ocupados pelos nós do plano de controle e serão responsáveis por um conjunto de códigos de dispersão. Estes nós armazenarão as tuplas  $\langle \text{Código}, ID \rangle$  nas quais o *Código* identifica um recurso físico

e o *ID* identifica o nó, do plano de controle, responsável pelo recurso em questão. A rede de informação, por sua vez, é sobreposta à rede do plano de controle e construída através de conexões TCP. Enquanto na rede do plano de informação são trocados pacotes TCP, na rede de controle são trocados pacotes UDP (*User Datagram Protocol*)

O plano de informação é construído de acordo com a estrutura semântica das aplicações P2P, descritas na seção 3.2.4, para permitir o acesso aos recursos da rede. Este plano é constituído de um espaço lógico, que compreende todos os códigos de dispersão e é dividido entre os nós do plano de informação.

A estrutura utilizada neste trabalho baseou-se no modelo CAN. Esta topologia foi escolhida devido ao fato de que a tabela de roteamento de cada nó não aumentar de acordo com o aumento da rede e sua semelhança com a topologia também desenvolvida por nós [65] possibilitando uma redução ainda maior do número médio de saltos entre os nós de forma determinística.

As principais atividades deste sistema são:

- **Registro de nó:** Para se registrar no plano de informação, um nó  $\alpha$  do plano de controle encontra, através do LMP, um vizinho ( $\beta$ ) no plano de controle que já esteja registrado no plano de informação. O nó  $\alpha$  envia ao nó  $\beta$ , pela rede do plano de controle, uma requisição de registro no plano de informação contendo seu *Código* e *ID*. A requisição é roteada na rede do plano de informação, através de  $\beta$  até atingir o nó do plano de informação responsável pelo código presente na requisição. Este nó é ocupado por um nó  $\gamma$  que responde ao nó  $\alpha$ , através da rede do plano de controle, informações pertinentes aos códigos que serão de responsabilidade do nó  $\alpha$  e relações de vizinhanças a serem estabelecidas. O nó  $\alpha$  envia, através da rede do plano de controle, a confirmação de recebimento das informações. A Figura 4.2 ilustra esta ativi-

dade sendo que as linhas cheias representam tráfego no plano de controle e as linhas tracejadas, tráfego no plano de informação. Esta notação é utilizada na ilustração de todas as atividades a seguir.

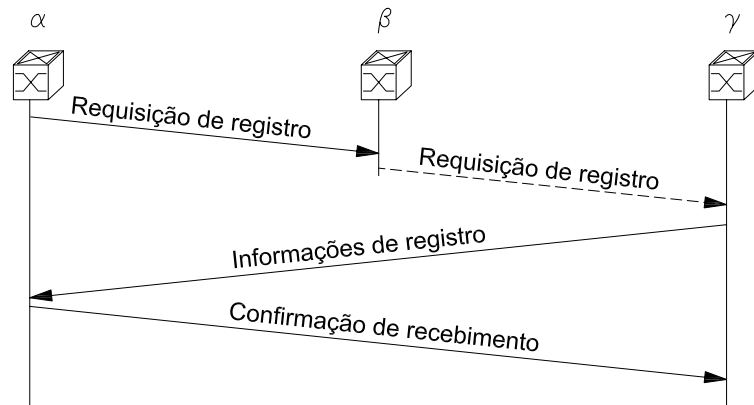


Figura 4.2: Atividades de registro.

- **Estabelecimento de vizinhança:** De posse dos *ID* dos nós com os quais deve ser estabelecer a relação de vizinhança, três mensagens são trocadas com cada vizinho para estabelecer a sessão TCP (SYN, SYN+ACK, ACK). Estas mensagens são trocadas através da rede de controle e esta atividade está ilustrada na Figura 4.3.
- **Publicação de Informações:** Para publicar um recurso, um nó roteia a requisição de publicação, contendo o *Código* do recurso e seu *ID*, através da rede do plano de informação. A requisição é roteada até que o nó do plano de informação responsável pelo *Código* da requisição seja encontrado. O nó encontrado armazena a tupla  $\langle \text{Código}, \text{ID} \rangle$  e envia, através da rede de controle, a confirmação de publicação. Os pacotes trocados estão apresentados

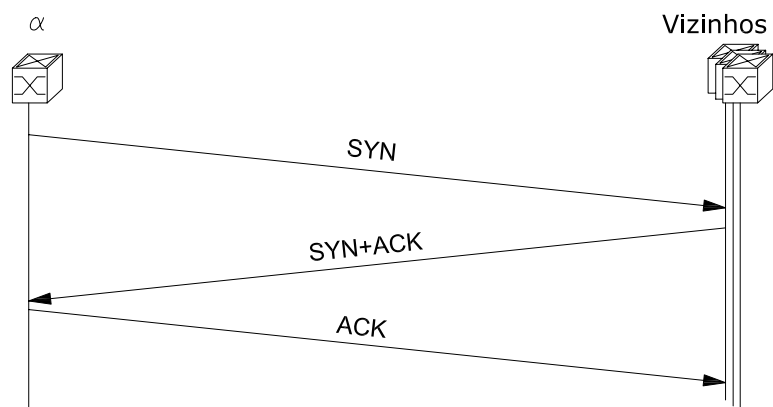


Figura 4.3: Atividades de estabelecimento de vizinhança.

na Figura 4.4.

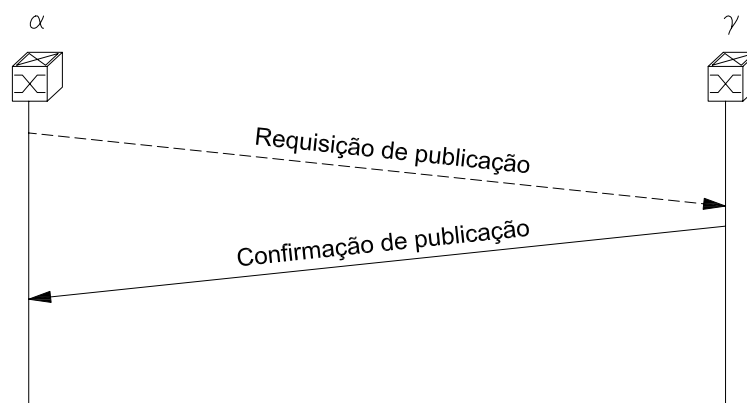


Figura 4.4: Atividades para publicação de informação.

- **Busca por Informações:** A requisição de busca é composta pelo *Código*

do recurso buscado e o *ID* do nó que requer a informação. Esta requisição é roteada através da rede de informação até encontrar o nó do plano de informação responsável pelo *código*. Este nó responde pela rede do plano de controle o *ID* armazenado para o *Código* buscado. Note que a resposta será nula se o recurso não existir ou não estiver disponível. Uma confirmação de recebimento é enviada de volta através da rede do plano de controle. Todo processo de busca pode ser visto na Figura 4.5.

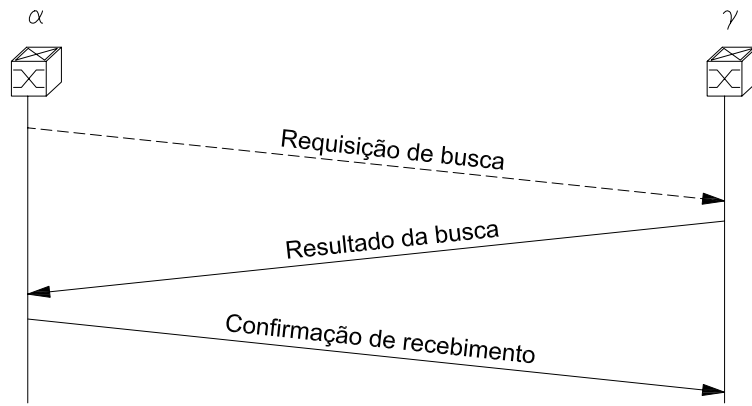


Figura 4.5: Atividades de busca de informação.

## 4.2 Exemplo de Construção do Plano de Informação

Um plano de informação com um número fixo de nós e com relações de vizinhanças já pré-estabelecidas é assumido. Os nós da rede física terão um código de dispersão gerados a partir da função de dispersão, que definirá em qual nó virtual esse deve se registrar. Após o registro estes dividem a gerência dos nós entre si e definem uma relação de vizinhança estabelecendo conexões através do TCP.

Para ilustrar a operação do sistema utilizaremos as redes da Figura 4.6. Na Figura 4.6(a) temos a rede física, também adotada com rede do plano de controle, e na Figura 4.6(b) a rede utilizada no plano de informação. Não esquecer que a rede do plano de informação tem topologia toroidal representada aqui de forma bi-dimensional. Cada nó da rede do plano de informação armazena 16 posições da tabela de dispersão e este sistema comportaria um número máximo de 256 recursos como mostrado na Figura 4.6(c).

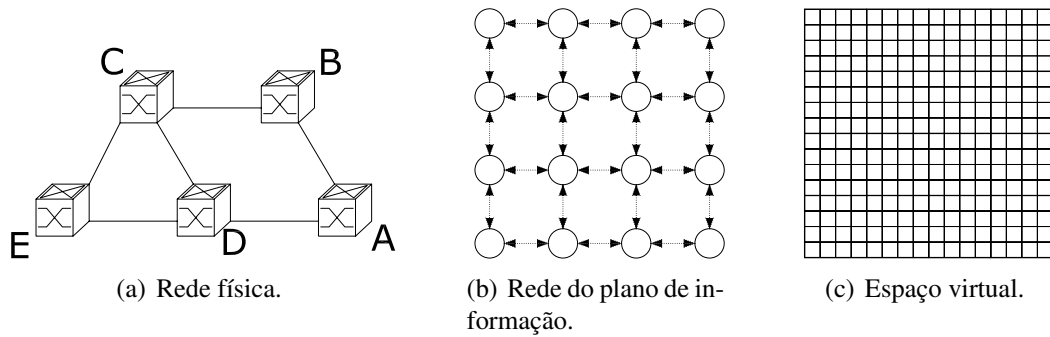


Figura 4.6: Redes adotadas.

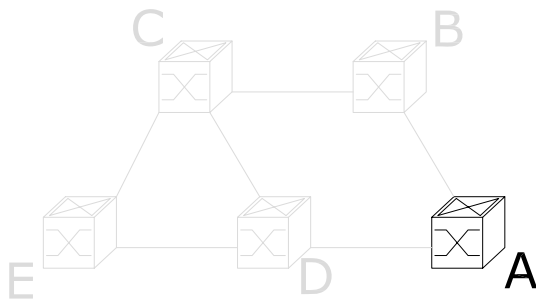
Inicialmente ilustraremos o sistema com apenas um nó físico, conforme ilustrado na Figura 4.7(a). Este exemplo ilustra o fato de o primeiro nó físico (*A*) tornar-se disponível na rede. Este nó gera seu próprio código de dispersão, que juntamente com o código dos outros nós, está ilustrado na Tabela 4.1. Ao perceber, através do protocolo LMP, que não existe nenhum vizinho ativo, o nó *A* gera seu próprio plano de informação, registrando-se no ponto indicado pelo código gerado. Na Figura 4.7(b) é mostrado que este nó físico ocupa todos os nós do plano de informação e nó do plano de informação em que o nó físico *A* foi registrado é representado em negrito.

Com a entrada do nó *B*, formando a rede física ilustrada na Figura 4.8(a), os procedimentos iniciais até percepção de um vizinho são realizados. A partir de

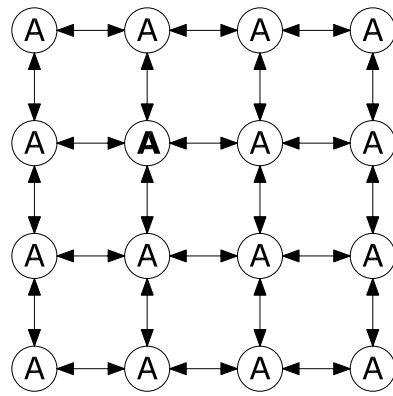


Tabela 4.1: Código dos Nós

Nó	Código
Nó A	(07,11)
Nó B	(14,15)
Nó C	(13,03)
Nó D	(13,12)
Nó E	(03,11)



(a) Rede física com apenas um nó.



(b) Plano de informação correspondente a rede física com apenas um nó.

Figura 4.7: Alterações na redes: Entrada do primeiro nó.

então os seguintes pacotes são trocados:

- o nó *B* envia requisição de registro, através da rede do plano de controle, ao nó *A* contendo o *Código* gerado e seu *ID*;
- o nó *A*, ao perceber que o *Código* de *B* está em um nó ocupado por ele divide o espaço virtual, conseqüentemente, os códigos a serem administrados, envia para *B*, pela rede do plano de controle, uma lista de códigos a serem administrados por este e o *ID* dos vizinhos a serem estabelecidos (*ID* de *A*);

- o nó *B* confirma o recebimento das informações enviando um pacote ao nó *A* através da rede do plano de controle;
- o nó *B* inicia o procedimento de estabelecimento de vizinhança, enviando ao nó *A* através da rede física um pacote SYN para iniciar uma sessão TCP;
- o nó *A* envia o SYN+ACK pelo mesmo meio;
- o nó *B* estabelece a sessão através do envio de um ACK ao nó *A* através mesmo meio novamente.

A configuração resultante da rede do plano de informação é ilustrada na Figura 4.8(b).

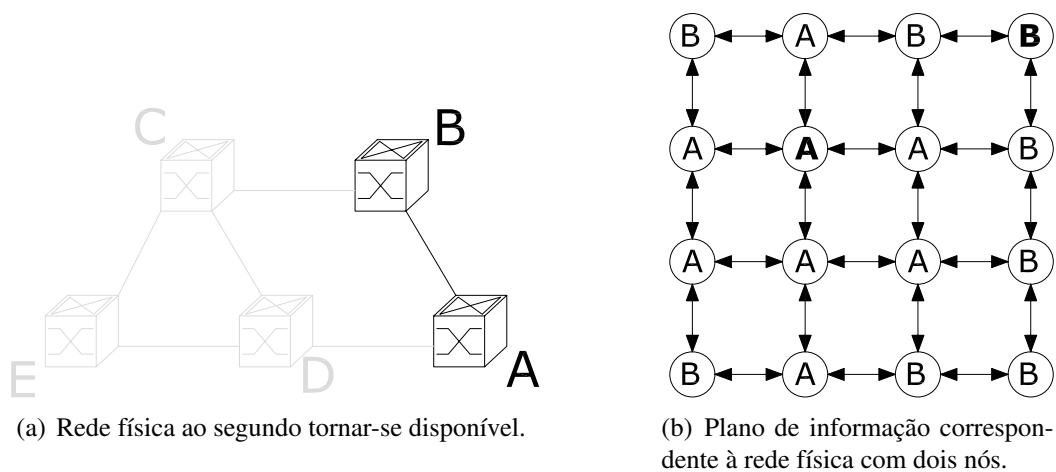


Figura 4.8: Alterações na redes: Entrada do segundo nó.

Um terceiro nó (*C*) ao entrar no sistema, ilustrado na Figura 4.9(a), repete os procedimentos de *B*. Caso a requisição seja enviada ao nó que não seja responsável pelo código de dispersão da requisição, o nó que a recebeu, a envia para o vizinho mais próximo do ponto indicado pelo *Código*. Este envio é feito através da rede do plano de informação até que a requisição atinja o nó responsável pelo seu código. A

partir de então, os procedimentos de registro, divisão dos nós do plano de informação, envio de tabela de dispersão (correspondente ao intervalo de gerência definido) e definição de vizinhança se repetem resultando no plano de informação ilustrado na Figura 4.9(b).

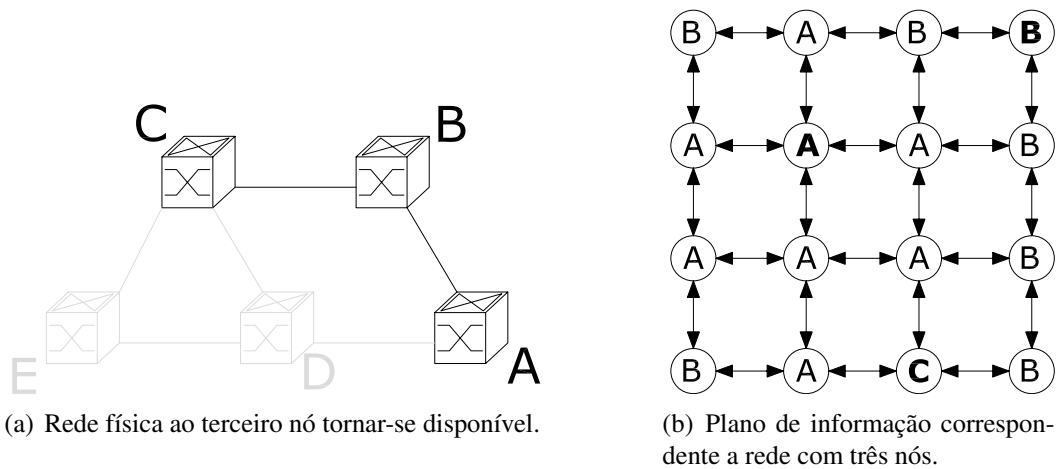


Figura 4.9: Alterações na redes: Entrada do terceiro nó.

Este procedimento se repete para todos os outros nós ingressantes no sistema conforme ilustrado na Figura 4.10(a). Caso um nó, ao entrar no sistema, tiver seu código em um nó do plano de informação em que já tenha um nó físico, o novo nó é registrado, porém, este não altera o plano de informação e não gerencia nó deste plano, como é o caso do nó *D*. O plano de informação correspondente à rede física com todos os nós é ilustrada na Figura 4.10(b). A Figura 4.11 ilustra a rede do plano de informação formada através do TCP. Note que esta não é uma rede em malha completa.

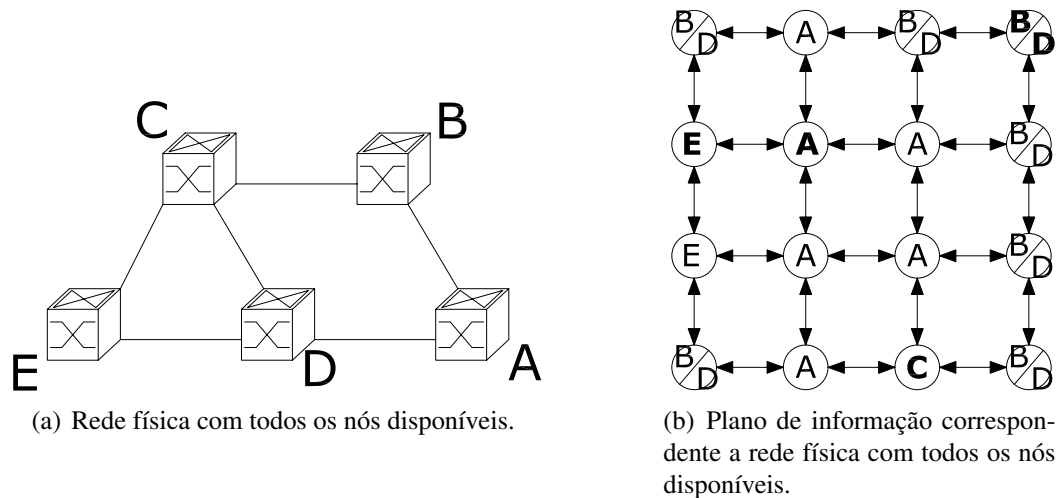


Figura 4.10: Alterações na redes: Rede com todos os nós.

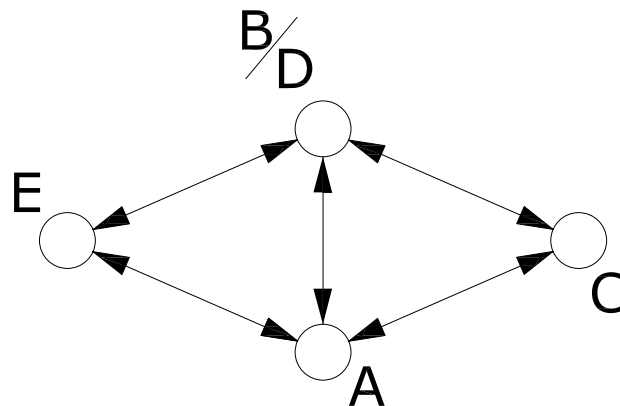


Figura 4.11: Rede do plano de informação formada pelas relações de vizinhanças através do TCP.

### 4.3 Exemplo de Acesso às Informações

Uma vez registrado, um nó pode iniciar a publicação dos recursos por ele gerenciado. Este nó aplica a função de dispersão a cada um dos seus recursos e juntamente com seu identificador (endereço MAC ou IP) envia uma requisição de publicação no plano de informação. Tal requisição é roteada através do plano

de informação até que o nó responsável por aquele código de dispersão seja encontrado. Então, o código, é inserido na tabela de dispersão deste nó, na posição correspondente em conjunto com o identificador do nó que gerencia o recurso em questão.

Para ter acesso ao identificador de um nó responsável por um recurso desejado deve-se proceder de forma semelhante à publicação. Gera-se uma requisição de identificador com o código hash obtido aplicando-se a função de dispersão ao recurso desejado e o identificador do autor da requisição. Esta é roteada pela rede do plano de informação até encontrar o nó físico responsável por armazenar o código buscado. Este nó por sua vez, responde ao autor da requisição o endereço do nó físico responsável pelo recurso. Desta forma, torna-se possível contatar o nó responsável pelo recurso através do protocolo RSVP-TE e reservá-lo.

Para ilustrar o processo será feita a publicação dos recursos do nó  $E$  mostrados na Tabela 4.2.

Tabela 4.2: Código dos Recursos

Item	Recursos	Código
1	Comprimento de Onda $\lambda_1$ , E-B, 10 Gbps	(01,05)
2	Comprimento de Onda $\lambda_2$ , E-C, 2,5 Gbps	(10,00)

Para o recurso 1, ao gerar a requisição de publicação, o nó  $E$  percebe que ele próprio é responsável pelo nó onde o recurso será publicado, então acrescenta em sua tabela de dispersão na posição (01,05) seu próprio identificador. Já para o recurso 2 ele percebe que o ponto buscado não está sob sua gerência e então passa a requisição ao nó gerenciado pelo nó  $B$ . Este por sua vez passa para um nó do plano de informação gerenciado pelo nó  $C$  que insere em sua tabela de dispersão na posição (10,0) o identificador do nó  $B$ . Uma mensagem confirmando a publicação

é enviada ao nó  $B$  através da rede do plano de controle.

A busca é realizada de maneira semelhante, diferenciando pela resposta, através da rede de controle, ao nó que originou a requisição do identificador armazenado na posição desejada.

## 4.4 Exemplo de Estabelecimento do Caminho Óptico

Para ilustrar, vamos supor um caminho fim-a-fim entre dois nós, por exemplo  $A$  e  $C$  no comprimento de onda  $\lambda_1$  e taxa de 10Gbps. O nó de origem deve buscar recursos que satisfaçam as restrições impostas com origem em  $A$ . Apenas o enlace entre  $A - D$  possui o comprimento de onda  $\lambda_1$  com taxa disponível de 10 Gbps disponível com origem em  $A$ , porém várias buscas são realizadas, inclusive por informações de recursos que não estão disponíveis no momento ou simplesmente não existem. O recurso encontrado é reservado através do RSVP-TE. Porém o destino ainda não foi atingido, fazendo com que o nó  $A$  faça buscas por enlaces com origem no nó  $D$ . Novamente várias buscas são feitas e são encontrados os enlaces  $D - A$ ,  $D - C$  e  $D - E$  que possuem uma taxa de 10Gbps disponível no comprimento de onda  $\lambda_1$ . Após a reserva dos recursos pelo RSVP-TE, nota-se que um caminho foi encontrado e então se estabelece um caminho fim-a-fim utilizando os recursos escolhidos (comprimento de onda  $\lambda_1$  com 10 Gbps disponíveis nos enlaces entre os nós  $A-D$  e  $D-C$ ) e libera-se os recursos não utilizados. Os critérios de escolha dos recursos a serem utilizados no estabelecimento dos caminhos ópticos não fazem parte do escopo deste trabalho. Neste, nos limitamos apenas às questões pertinentes ao acesso à informação do estado da rede por seus nós.

## 4.5 Reconfiguração Determinística

O roteamento no plano de informação é uma tarefa fundamental para o bom funcionamento do sistema, e uma vez que a falha no roteamento indicará que um recurso buscado não existe. Também temos que destacar que o desempenho do roteamento é de extrema importância na redução do tráfego de controle das redes, umas das principais metas deste trabalho. No modelo proposto, o roteamento na rede do plano de informação dar-se-á pelo melhor caminho, uma vez que todas as ligações são pré-estabelecidas o que permite a construção de uma tabela de roteamento ótima.

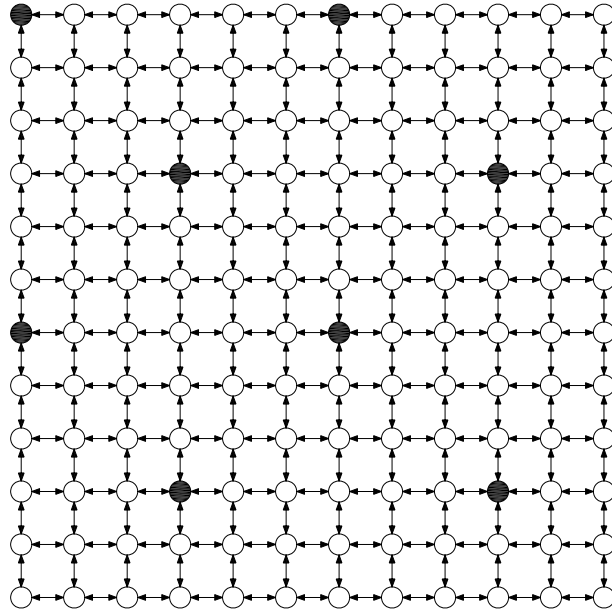
Em grandes planos de informação o processo de roteamento pode ser feito através de muitos saltos, ameaçando a eficiência do modelo. Para aumentar a eficácia do roteamento (reduzir o número de saltos médio entre os nós do plano de informação) adota-se reconfiguração determinística de topologia para obtenção de característica de Mundo Pequeno como descrito na Seção 3.3.

Esta reconfiguração consiste em eleger alguns nós  $CO$  para tornarem-se concentradores.  $CO$  tem valor determinado pela Equação 4.1 na qual  $n$  é o numero total de nós no plano de informação e  $d$  é a máxima distância entre um nó e um concentrador, também chamado de Fator de Mundo Pequeno (FMP). Sendo  $V$  o número médio de vizinhos por nó (incluído os de curta e longa distância) e  $g_o$  o números de vizinhos de curta distância os concentradores são interligados segundo a Equação 4.2 na qual o concentrador  $i$  conecta aos  $V - g$  nós  $j$ . Pode-se observar na Figura 4.12(a) uma rede virtual com 144 nós do plano de informação, onde 8 deles foram eleitos concentradores para  $d = 3$ . Na Figura 4.12(b) mostramos a interconexão dos concentradores através de conexões TCP. Um número médio de vizinhos ótimo,  $V$  é apresentado em [66] e também será calculado neste trabalho posteriormente.

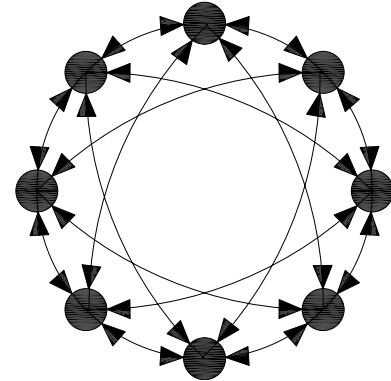
$$CO = 2 \times \left\lceil \sqrt{n}/2d \right\rceil^2 \quad (4.1)$$

$$j = \text{mod} \left( i + 1 + \left\lfloor k \times \frac{CO}{V - g_o} \right\rfloor ; CO \right), \quad k = 0, 1, 2, \dots, (V - g_o - 1) \quad (4.2)$$

$$j, i \in \{0, 1, 2, \dots, CO - 1\}$$



(a) Eleição dos Concentradores (nós preenchidos).



(b) Rede entre concentradores.

Figura 4.12: Reconfiguração determinística.

Com esta reconfiguração espera-se reduzir o tráfego na rede de controle do sistema e o tempo de roteamento de uma requisição.



## 4.6 Sumário

A proposta de construir uma plano de informação, apresentada neste Capítulo, figura pela primeira vez na literatura. Através do mecanismo proposto conseguimos distribuir informações sobre recursos físicos em redes multi-camadas de forma estruturada, o que possibilita a eliminação do mecanismo de distribuição de informação por inundação. Um exemplo dos procedimentos necessários ao funcionamento dos enlaces é descrito e uma proposta de reconfiguração determinística, da topologia utilizada, é proposta.

# Capítulo 5

## Estudo Analítico

Uma abordagem analítica será feita neste Capítulo com o objetivo de mapear a influência dos parâmetros do sistema em seu desempenho. O número de pacotes por enlace e o número de saltos necessários para realizar as atividades descritas anteriormente serão analisados tanto para o modelo proposto como para o modelo clássico baseado em inundação.

### 5.1 Considerações Iniciais

Os benefícios e limitações do plano de controle baseado em DHT como meio de distribuição de informação em planos de controle podem ser facilmente avaliados através de simuladores como NS2 [67]. Simuladores numéricos podem utilizar modelos detalhados de redes para protocolos e componentes e, os resultados são avaliações confiáveis do desempenho. Contudo, estes simuladores demandam simultaneamente tempo e capacidade de processamento cada vez maiores à medida em que a rede em estudo aumenta. Além disso, a complexidade de uma rede multi-camadas, quando analisados em uma grande variedade de parâmetros, transformam

a interpretação dos resultados e a determinação da influência de cada parâmetro em uma tarefa muito complexa.

Uma abordagem analítica é proposta neste trabalho utilizando-se de topologias físicas regulares e avaliando procedimentos essenciais para instalação e operação do plano de controle baseado em DHTs. Estas considerações permitem um estudo analítico do número de saltos médios na rede física, então o número de saltos na rede virtual pode ser convertido em saltos na rede física. Portanto, a avaliação do desempenho ocorrerá sobre as seguintes métricas:

- Carga por enlace: A quantidade de pacotes trocados pelo mecanismo de distribuição de informação utilizando recursos físicos que poderiam ser utilizados para transmissão de dados. Melhorias no desempenho do sistema quanto a essa métrica é o principal objetivo desta proposta.
- Atraso: Número de saltos não simultâneos, utilizados para disponibilizar e acessar informações. Traduz atrasos de propagação e espera em filas em um percurso. Portanto, o número de saltos é um dado importante para se avaliar o atraso no acesso a informações.

## 5.2 Atividades Analisadas

Os procedimentos citados no Capítulo 4 podem ser divididos em tráfego de instalação e de operação. Esta divisão permite avaliar os custos fixos, independentes do uso da rede, e as variáveis, que dependem do uso da rede. Em outras palavras, o tráfego de instalação será formado pelos pacotes trocados para estabelecer o sistema antes do início dos procedimentos de busca e atualização e o tráfego de operação é o utilizado no atendimento às demandas dinâmicas de tráfego. Portanto, o tráfego de

instalação é composto pelo: registro dos nós, definição da vizinhança e publicação de recursos. O tráfego de operação envolve a busca e atualização de informações.

### 5.3 Metodologia

Para avaliação de desempenho assumiremos uma função de dispersão sem conflitos e com dispersão uniforme das chaves a serem publicadas. A rede do plano de informação possuirá topologia fixa, com a mesma quantidade de nós da rede do plano de controle, permitindo assim comparação direta com o modelo baseado em inundação. Será considerado uma rede física sem perdas de topologia regular e o número médio de saltos será considerado sempre pelo menor caminho. As atividades envolvendo busca, publicação e atualização geram uma transação por informação, isto é, não há agrupamento de vários parâmetros em apenas um LSA.

O tráfego gerado pelo sistema será avaliado através do número médio de pacotes por enlace físico. Esta métrica pode indicar a porcentagem de uso da capacidade dos transceptores utilizada pelo sistema de distribuição e busca de informações.

O atraso será avaliado através do número de saltos na rede física para se executar a atividade em questão. Portanto, para publicação de todos os recursos, o número de saltos avaliados será o do caminho crítico e para busca e atualização de informações o caminho médio será considerado. Note que esta métrica pode indicar quanto tempo uma informação desatualizada pode permanecer disponível.

O mecanismo de inundação para disseminação de informação será sempre utilizado como parâmetro de comparação, uma vez que o presente trabalho propõe uma alternativa a este mecanismo. O tráfego de instalação é composto pela publicação dos recursos que se dará através da inundação na rede. Da mesma maneira, a operação será composta pela atividade de atualização. Por simplicidade, será con-

siderado apenas uma mensagem de publicação e uma confirmando o recebimento em cada enlace da rede durante essas operações como acontece em redes lineares. Esta decisão foi tomada visando obter o melhor caso de ocupação de enlaces pelo modelo baseado em inundação.

Estas estimativas foram feitas em função das seguintes variáveis:

- $S_{RF}$  – Número médio de saltos nos caminhos entre um nó e todos os outros da rede física. Seu valor depende do número de nós da rede física ( $n$ ) e da topologia adotada.
- $S_{RV}$  – Número médio de saltos nos caminhos entre um nó e todos os outros da rede virtual. Para uma rede virtual sem a reconfiguração determinística de topologia  $S_{RV}$  é dado pela Equação 5.1 e com a reconfiguração determinística de topologia  $S_{RV}$  é dado pela Equação 5.2.

$$S_{RV} = \left\lceil \frac{\sqrt{n}}{g_o/2} \right\rceil \quad (5.1)$$

$$S_{RV} = \left\lceil \frac{\left\lceil \frac{2d+1}{3} \right\rceil + \left( \left\lceil \frac{C}{V-g_o} \right\rceil + 2 \left\lceil \frac{2d+1}{3} \right\rceil \right) (C-1)}{C} \right\rceil \quad (5.2)$$

- $N_E$  – Número total de enlaces na rede física, também dependente da topologia da rede e do número de nós.
- $V$  – Número médio do total de relações de vizinhança estabelecidas através do TCP.
- $g_o$  – Número de vizinhos de cada nó da rede do plano de informação mantém sem reconfiguração da topologia. Neste trabalho utilizamos sempre  $g_o = 4$ .

- $S_{RFC}$  – Número médio de saltos nos caminhos críticos de cada nó da rede física. Seu valor também depende do número de nós da rede física e da topologia adotada.
- $S_{RVC}$  – Número médio de saltos nos caminhos críticos de cada nó da rede física. Para uma rede virtual sem a reconfiguração determinística de topologia  $S_{RVC}$  é dado pela Equação 5.3 e com a reconfiguração determinística de topologia  $S_{RVC}$  é dado pela Equação 5.4.

$$S_{RVC} = \sqrt{n} - 1 \quad (5.3)$$

$$S_{RVC} = 2d + \frac{C}{V - g_o} \quad (5.4)$$

## 5.4 Equacionamento

Nesta seção será apresentado o equacionamento, da carga por enlace e do atraso, durante as atividades desenvolvidas pelo mecanismo de disseminação e busca de informação. As atividades são divididas em atividades de instalação e de operação.

Para possibilitar uma comparação direta também serão sucintamente apresentadas as equações da quantidade de pacotes por enlace e do número de saltos não simultâneos, de instalação e operação, do sistema completamente baseado em inundação [68].

### 5.4.1 Atividade de Instalação

- **Registro dos Nós:** Ao se registrar, um nó envia uma requisição de registro a um vizinho no plano de informação. Então a requisição é roteada no plano

de informação até atingir o destino (cada salto no plano de informação são vários saltos na rede física). O nó responsável pela posição envia uma confirmação (ACK) para o nó que originou a requisição. A quantidade de pacotes gerada por enlace  $CIRN$  é dada pela Equação 5.5 e o número de saltos não simultâneos  $AIRN$  é dado na Equação 5.6.

$$CIRN = \frac{1 + S_{RF}S_{RV} + S_{RF}}{N_E} \quad (5.5)$$

$$AIRN = 1 + S_{RF}S_{RV} + S_{RF} \quad (5.6)$$

- **Estabelecimento da Vizinhança:** O nó troca 3 pacotes com cada um dos  $V$  vizinhos para estabelecer uma sessão TCP (SYN, SYN+ACK, ACK). A quantidade de pacotes por enlace  $CIEV$  é dado pela equação 5.7 e o número de saltos  $AIEV$  dado pela equação 5.8.

$$CIEV = \frac{(S_{RF} + S_{RF} + S_{RF}) \times V}{N_E} \quad (5.7)$$

$$AIEV = S_{RF} + S_{RF} + S_{RF} \quad (5.8)$$

- **Publicação de Recursos:** O nó roteia pelo plano de informação uma requisição de publicação para cada recurso ( $RP$ ) até encontrar seus pontos de publicação. Após a inserção dos identificadores na tabela de dispersão os nós enviam uma confirmação (ACK). A quantidade de pacotes por enlace  $CIPR$  é dado pela equação 5.9 e o número de saltos  $AIPR$  dado pela equação 5.10.

$$CIPR = RP \times \frac{S_{RF}S_{RV} + S_{RF}}{N_E} \quad (5.9)$$

$$AIPR = S_{RVC} \times S_{RF} + S_{RF} \quad (5.10)$$

Portanto, temos a quantidade de pacotes total durante os procedimentos de instalação ( $C_I$ ) dada pela soma das Equações 5.5, 5.7 e 5.9 apresentada na Equação 5.11 e o número total de saltos não consecutivos durante a instalação ( $A_I$ ) dada pela soma das Equações 5.6, 5.8 e 5.10 apresentada na Equação 5.12.

$$C_I = \frac{1 + S_{RF} \times [S_{RV} (1 + RP) + 1 + 3V + RP]}{N_E} \quad (5.11)$$

$$A_I = 1 + S_{RF} \times (S_{RVC} + S_{RV} + 5) \quad (5.12)$$

Para os sistemas que disponibilizam as informações através de inundação a quantidade de pacotes por enlace durante a instalação,  $C_{II}$ , se comporta conforme a Equação 5.13 e o número de saltos não consecutivos,  $A_{II}$ , dado pela Equação 5.14

$$C_{II} = 2RP \quad (5.13)$$

$$A_{II} = S_{RFC} \quad (5.14)$$

#### 5.4.2 Atividades de Operação

- **Busca por Recurso:** O nó roteia a requisição de informação pela rede até encontrar o código requerido. O nó responsável por este código responde ao



criador da requisição, o *ID* da informação desejada e recebe de volta uma confirmação (ACK). A quantidade de pacotes durante o processo *COBR* é dada pela Equação 5.15 e o número de saltos *AOBR* é dada pela Equação 5.16.

$$COBR = \frac{S_{RF}S_{RV} + S_{RF} + S_{RF}}{N_E} \quad (5.15)$$

$$AOBR = S_{RF}S_{RV} + S_{RF} \quad (5.16)$$

- **Atualização de Recurso:** O nó roteia a requisição de publicação pela rede até encontrar a posição requerida. O nó detentor da informação responde ao criador da requisição uma confirmação (ACK). A quantidade de pacotes durante o processo *COAR* é dada pela Equação 5.17 e o número de saltos *AOAR* é dada pela Equação 5.18.

$$COAR = \frac{S_{RF}S_{RV} + S_{RF}}{N_E} \quad (5.17)$$

$$AOAR = S_{RF} \times S_{RV} \quad (5.18)$$

Portanto, temos a quantidade total de pacotes, por enlace, trocados durante os processos de operação *C<sub>O</sub>*, dada pela soma das Equações 5.15 e 5.17 apresentada na Equação 5.19 e número total de saltos não simultâneos durante as atividades de operação, *A<sub>O</sub>*, dada pela soma das Equações 5.16 e 5.18 apresentada na Equação 5.20.

$$C_O = \frac{S_{RF} \times (2S_{RV} + 3)}{N_E} \quad (5.19)$$

$$A_O = S_{RF} \times (2S_{RV} + 1) \quad (5.20)$$

Para os sistemas que disponibilizam as informações através de inundação a quantidade de pacotes, por enlace, trocados durante as atividades de operação,  $C_{OI}$ , se comporta conforme a Equação 5.21 e o número total de saltos não consecutivos de operação,  $A_{OI}$ , se comporta conforme a Equação 5.22.

$$C_{OI} = 2 \quad (5.21)$$

$$A_{OI} = S_{RFC} \quad (5.22)$$

### 5.4.3 Número Ótimo de Vizinhos

O valor ótimo para o número médio de vizinhos de cada nó ( $V$ ) para se obter a menor quantidade de pacotes por enlace pode ser obtido através da minimização (por derivação) da soma das equações 5.11 e 5.19. O resultado é mostrado na equação 5.23.

$$V = g_v + \sqrt{\frac{2(C-1)(RP+3)}{3}} \quad (5.23)$$

## 5.5 Considerações Práticas

O equacionamento das métricas propostas para as atividades de operação não nos possibilita uma comparação fiel entre os dois modelos. Como descrito no Capítulo 4 o mecanismo proposto necessita de inúmeras buscas e pré-reservas para encontrar os recursos necessários ao estabelecimento de um enlace fim-a-fim. No modelo baseado em inundação, os dados já estão armazenados na LSBD não sendo necessário a busca por informação e as reservas acontecem depois do caminho definido e apenas nos recursos que serão utilizados. Porém, as informações armazenadas no LSBD podem estar desatualizadas e diferentes em cada nó, enquanto no modelo proposto, a busca por informação sempre que necessário reduz a possibilidade de se utilizar informações inconsistentes.

## 5.6 Sumário

Este Capítulo apresentou um estudo analítico simples do sistema de distribuição de informação proposto e do modelo clássico baseado em inundação. Este estudo possibilita, através de simulações numéricas, a avaliação da influência de cada parâmetro no desempenho do sistema. A extensão do estudo ao modelo clássico baseado em inundação possibilita a comparação dos desempenhos para as duas métricas descritas.

# Capítulo 6

## Análise dos Resultados

O objetivo deste capítulo é comparar o mecanismo estruturado de distribuição de informação com o mecanismo baseado em inundação amplamente utilizado nos planos de controle e validar os ganhos de desempenho obtidos com a reconfiguração determinística de topologia propostas nesse trabalho.

### 6.1 Metodologia

O resultados apresentados nos gráficos a seguir foram obtidos para os seguintes mecanismos de distribuição de informação: inundação; estruturado (utilizando o plano de informação sem reconfiguração determinística de topologia proposta); e estruturado com FMP assumindo os valores 6, 3 e 1.

Para verificar a influência da topologia da rede física nos mecanismos avaliados são adotadas redes com topologia em anel e grade. Embora, ambas as topologias adotem o mesmo grau lógico ( $g_f = 4$ ), a topologia em grade possui um menor número de saltos médio entre dois nós. Assim a influência da topologia física pode ser estudada.

## 6.2 Topologia Física em Anel

Para obtenção dos resultados desta seção foi utilizada a topologia apresentada na Figura 6.1. Os parâmetros utilizados foram: número de nós ( $n$ ) variando de 256 até 10000, simulando o crescimento das redes ópticas; o número de recursos publicados por cada nó variando de 10 até 1000, representando o aumento dos parâmetros e equipamentos gerenciáveis do plano de controle que necessitam publicar informações. Com a definição da topologia da rede física as variáveis descritas no Capítulo 5 podem ser definidas para esta topologia.

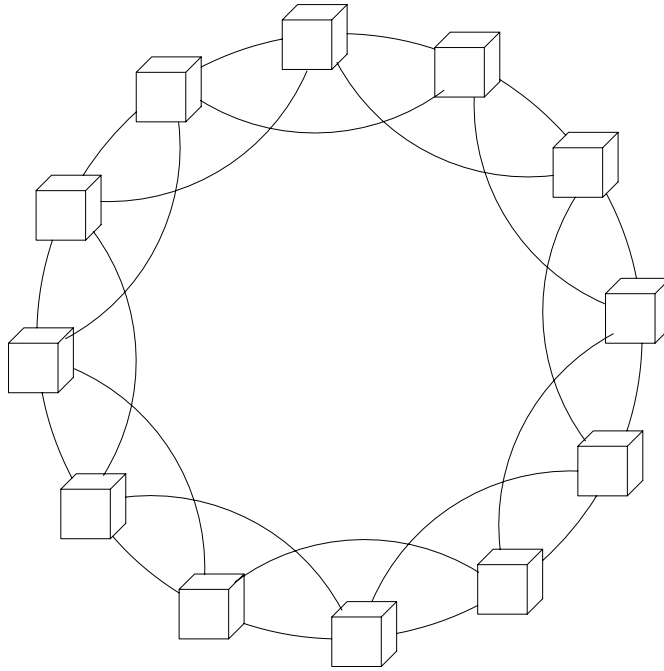


Figura 6.1: Topologia anel considerada na rede física

As Equações 6.1, 6.2 e 6.3 definem, respectivamente, o número médio de saltos nos caminhos entre um nó e todos os outros da rede física ( $S_{RF}$ ), o número total de enlaces na rede física ( $N_E$ ) e o número médio de saltos nos caminhos críticos de cada nó da rede física ( $S_{RFC}$ ).

$$S_{RF} = \frac{n}{g_f/2} \quad (6.1)$$

$$N_E = 2n \quad (6.2)$$

$$S_{RFC} = \frac{n}{g_f} \quad (6.3)$$

### 6.2.1 Carga por Enlace

A carga de instalação por enlace, isto é, a quantidade média de pacotes que precisa circular em cada enlace no momento em que o nó entra na rede é mostrado na Figura 6.2 para 128 recursos ( $RP$ ) a serem publicados por cada nó.

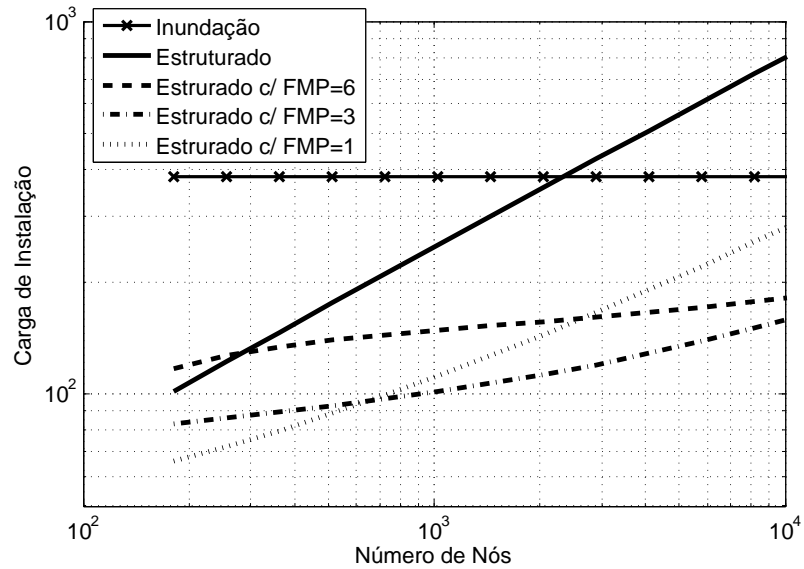


Figura 6.2: Carga de instalação por enlace em função do tamanho da rede.

Note que o modelo proposto com reconfiguração determinística apresenta resultados melhores para toda a faixa analisada. Todavia, o modelo estruturado apresenta

resultados piores para redes com mais de 2000 nós. Podemos ver também que a carga de instalação decresce com o aumento do FMP para grandes redes. Isto se deve ao fato de que o crescimento do número de nós, com um baixo FMP, implica em aumento expressivo na rede de concentradores levando ao aumento do número médio de saltos nesta rede. Os mecanismos propostos apresentam um crescimento da carga com o aumento do número de nós e somente o mecanismo estruturado com FPM=6 manterá um resultado melhor que o mecanismo baseado em inundação para redes muito maiores que 10000 nós.

Na Figura 6.3 a mesma carga de instalação é avaliada, porém o tamanho da rede ( $n$ ) foi fixado em 1024 nós e é variado o número de recursos a ser publicado, ou divulgado, inicialmente por cada nó ( $RP$ ) de 10 a 1000.

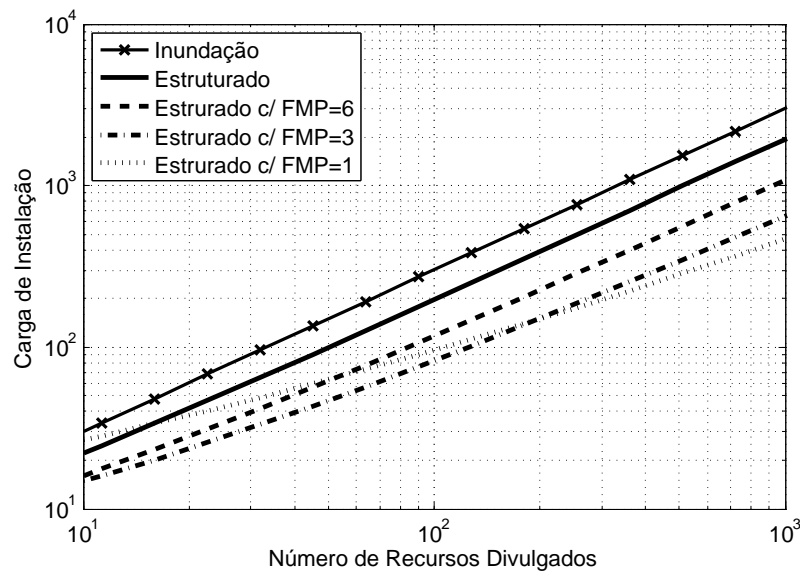


Figura 6.3: Carga de instalação por enlace em função da quantidade de informação a ser publicada.

Novamente, o modelo proposto apresentou desempenho melhor que o sistema baseado em inundação de quase uma ordem de grandeza quando 1000 informações

são publicadas. Porém, note que, inversamente ao comportamento da carga com variação do tamanho da rede, quanto menor o FMP, melhor o desempenho para grandes quantidades de recursos. Neste caso o número de nós permanece constante e o custo de registro e estabelecimento da vizinhança de vários concentradores é compensado com a redução do numero de saltos.

Na Figura 6.4 é avaliada a carga para se atualizar uma informação publicada, aqui chamada de carga de operação.

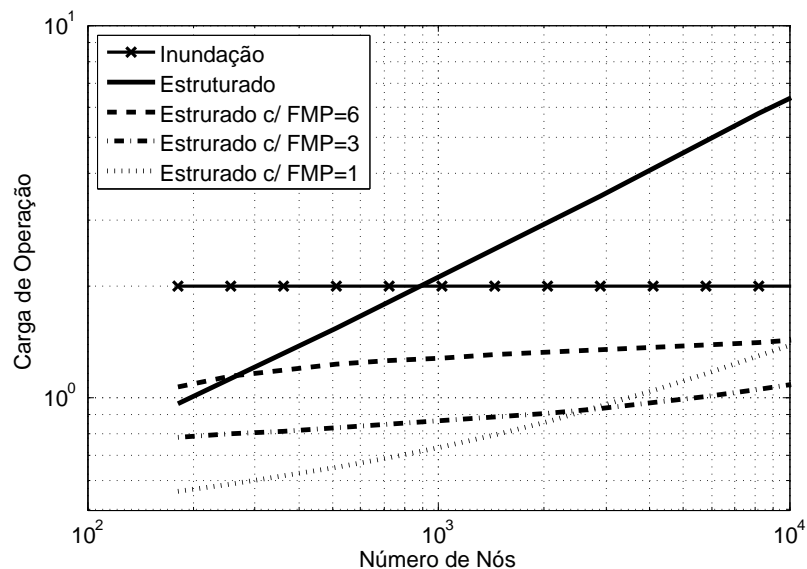


Figura 6.4: Carga de Operação por enlace em função da quantidade de recursos a ser publicada.

Resultados muito semelhantes aos apresentados para carga de instalação variando-se o número de nós, são obtidos. Somente o mecanismo proposto estruturado simples apresenta desempenho pior ao mecanismo baseado em inundação em certos tamanhos de rede.



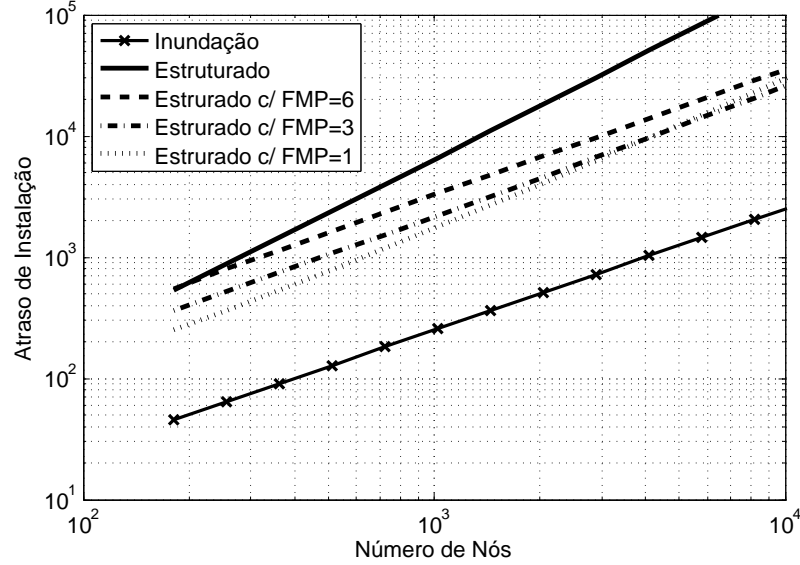


Figura 6.5: Atraso para instalação do sistema.

### 6.2.2 Número de Saltos

As avaliações dos atrasos para a rede em anel são apresentadas nas Figuras 6.5 e 6.6, para o atraso de instalação e operação respectivamente. Para análise do atraso de operação apenas a atualização é considerada.

Conforme esperado, o desempenho do sistema proposto em relação ao atraso apresenta-se pior. Porém a reconfiguração determinística tem papel importante, reduzindo pela metade a diferença do atraso do modelo estruturado simples e o baseado em inundação.

## 6.3 Topologia física em Grade

A topologia em grade adotada está apresentada na Figura 6.7 onde adotamos os mesmos intervalos para o número de nós e de recursos publicados que na topologia em anel. As equações 6.1 a 6.3 são redefinidas nas equações 6.4 a 6.6.

$$S_{RF} = \sqrt{n} - 1 \quad (6.4)$$

$$N_E = 2n - 2\sqrt{n} \quad (6.5)$$

$$S_{RFC} = 2\sqrt{n} - 2 \quad (6.6)$$

### 6.3.1 Carga por Enlace

A carga de instalação por enlace do novo sistema é mostrada na Figura 6.8 para 128 recursos publicados por cada nó.

O desempenho do sistema proposto para o tráfego de instalação variando o tamanho da rede foi novamente melhor que o baseado em inundação. Os mecanismos baseados em inundação apresentam uma carga constante por enlace, o que

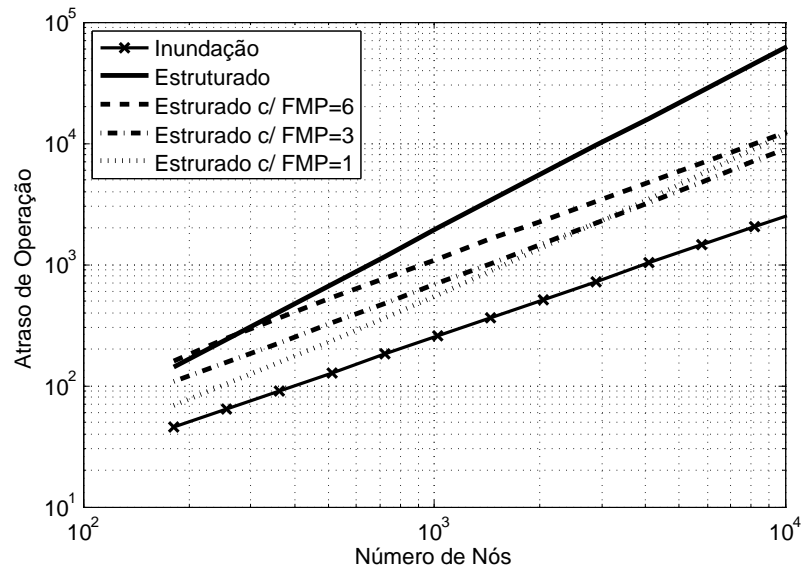


Figura 6.6: Atraso na operação do sistema.

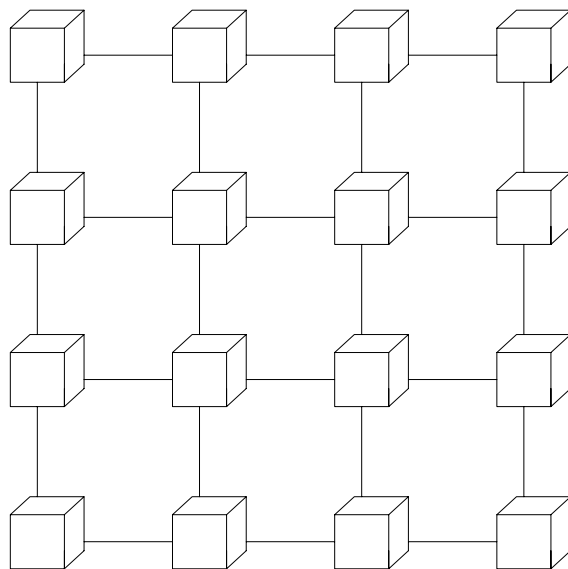


Figura 6.7: Topologia grade considerada na rede física

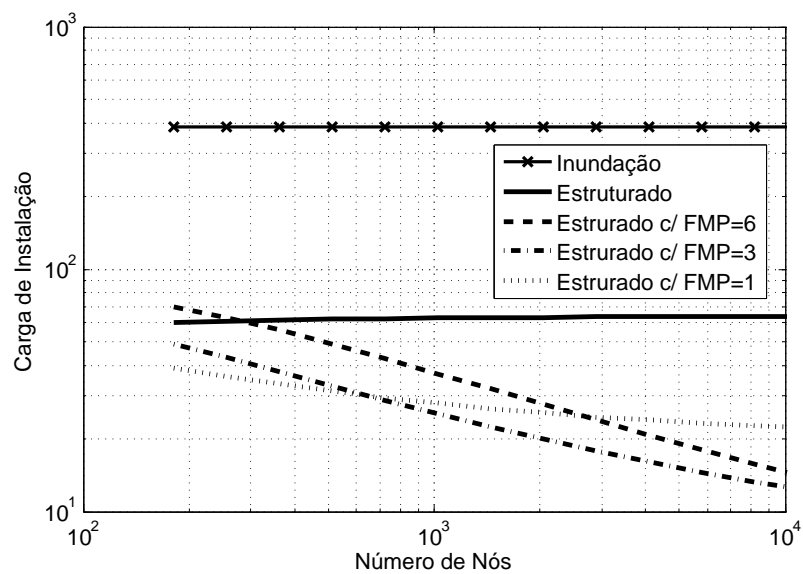


Figura 6.8: Carga de instalação por enlace em função do tamanho da rede.

também é apresentado pelo mecanismo estruturado simples, porém o mecanismo com reconfiguração determinística apresenta carga por enlace decrescente com o aumento do número de nós, característica crucial à escalabilidade do sistema. Isto

deve-se ao fato de a redução do número de saltos com a reconfiguração ser maior que o aumento do número de enlaces, sendo assim a carga fica menor em média em cada enlace.

Avaliando novamente o tráfego de instalação com um número fixo de nós ( $n = 1024$ ) e variando o número de recursos a serem publicados ( $RP$ ) de 10 a 1000 temos os resultados mostrados na Figura 6.9.

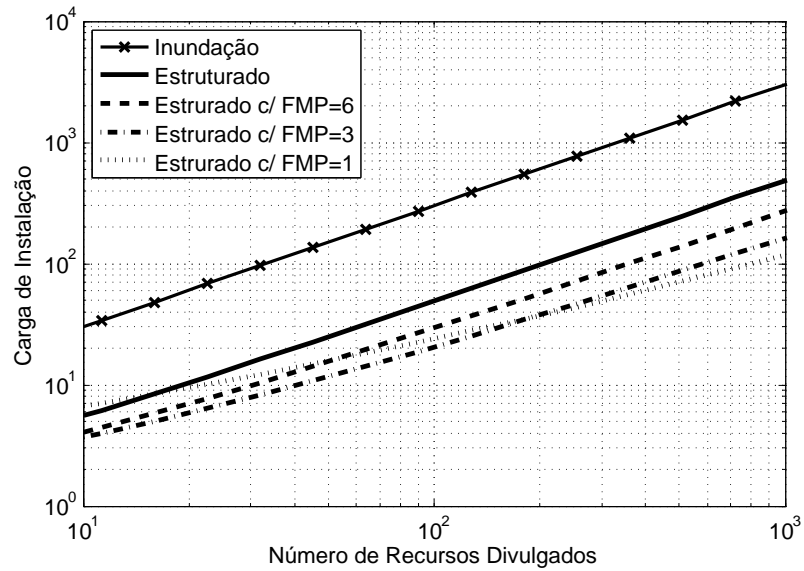


Figura 6.9: Carga de instalação por enlace em função da quantidade de recursos a ser publicada.

Temos uma diferença de quase uma ordem de grandeza quando 1000 informações são publicadas. Porém, note que, inversamente ao comportamento da carga com variação do tamanho da rede, quanto menor o FMP melhor o desempenho para grandes quantidades de recursos.

Para avaliarmos o tráfego de operação utilizamos o cenário no qual uma informação necessita ser atualizada. Na Figura 6.10 mostramos o desempenho para  $RP = 128$  e variando o número de nós da rede.

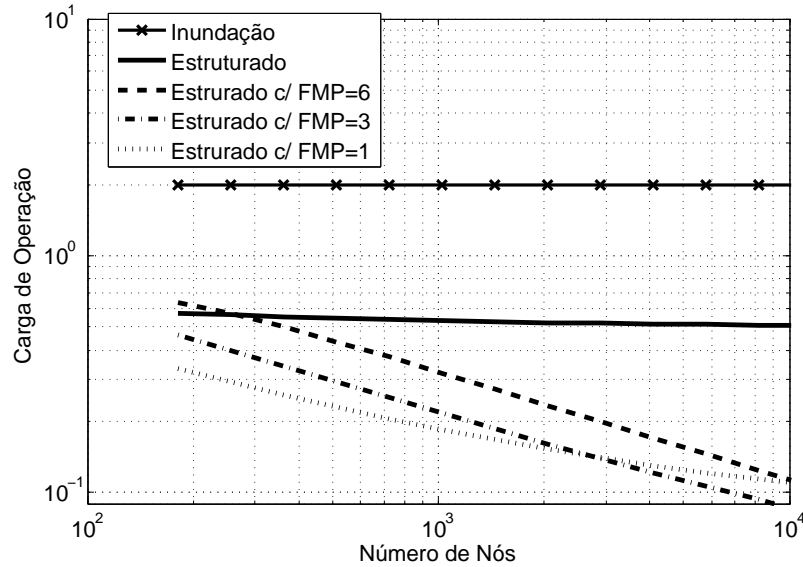


Figura 6.10: Carga de Operação por enlace em função da quantidade de recurso a ser publicada.

Resultados semelhantes aos apresentados na análise do tráfego de instalação porém deslocados devido ser apenas um recurso. A ocupação média por enlace cai com o crescimento da rede em todas as arquiteturas estruturadas, o que não é observado para a arquitetura baseada em inundação. Note também que atingimos com a arquitetura proposta, em média, menos de um pacote por enlace só possível devido ao mecanismo estruturado de distribuição de informação.

### 6.3.2 Número de saltos

A avaliação do número de saltos para a rede em grade são apresentadas nos dois gráficos seguintes. No primeiro é apresentado o número de saltos de instalação 6.11 e no segundo o de operação 6.12, lembrando que a operação considerada aqui se restringe apenas à atualização de informações da mesma forma que foi considerado para o estudo da carga. Devido ao fato de que a arquitetura baseada em

inundação sempre encontrará o menor caminho físico entre os nós é de se esperar que o desempenho do modelo proposto seja inferior.

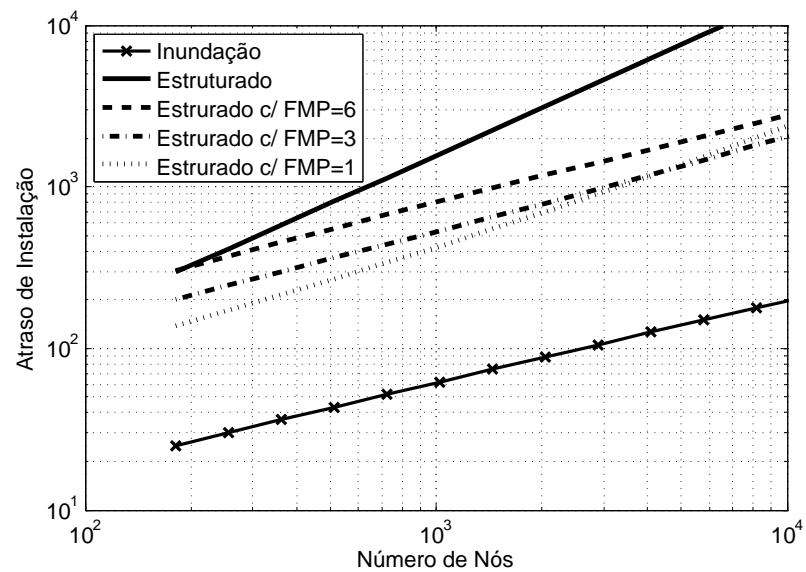


Figura 6.11: Atraso para instalação do sistema.

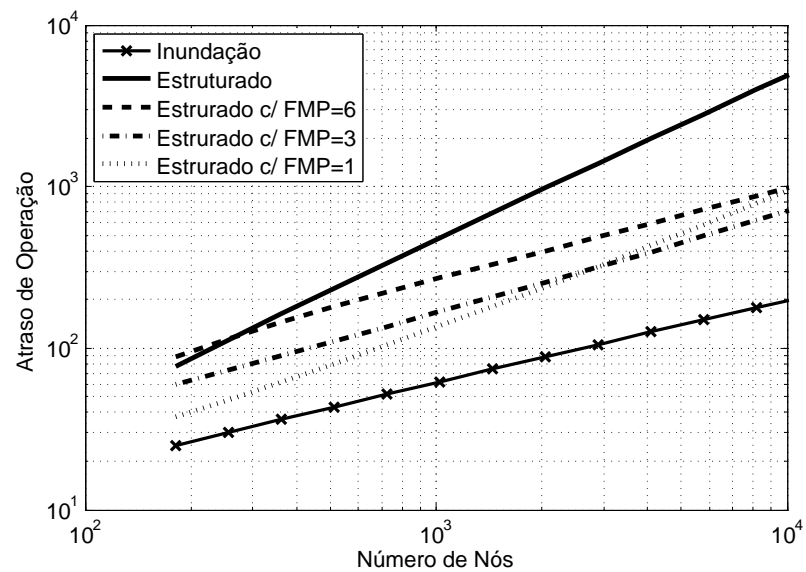


Figura 6.12: Atraso na operação do sistema.

Ambas figuras apresentam resultados semelhantes aos da rede em anel e desfavoráveis em mecanismos estruturados. Porém, a reconfiguração determinística produziu melhorias, e à medida que aumenta o tamanho da rede, diminui a diferença de desempenho.

## 6.4 Sumário

A comparação do desempenho da arquitetura proposta e da arquitetura baseada em inundação mostrou que uma redução considerável na taxa de ocupação dos enlaces é possível utilizando sistemas estruturados. Sistemas estruturados com reconfiguração determinística de topologia mostraram uma redução ainda maior. A penalidade do aumento no número de saltos está presente no modelo estruturado, porém, a penalidade é atenuada com o uso da reconfiguração determinística de topologia proposta.

## Capítulo 7

### Conclusão e Perspectivas

A evolução da infra-estrutura de transporte das redes de comunicação segue para um modelo formado por roteadores de alta velocidade interconectados por fibras ópticas. As redes ópticas, por sua vez, apresentam-se cada vez mais flexíveis com elementos ativos que necessitam ser gerenciados. Esta mudança aumenta ainda mais a importância da gerência neste sistema para possibilitar o atendimento das dinâmicas demandas de tráfego.

Contudo, o mecanismo de distribuição de informação adotado nas redes multicamadas é ineficiente. Além disso, medidas como redução da frequência de divulgação e restrição da inundação (criação de ASs) têm sido tomadas para reduzir o número de pacotes de controle na rede.

Aplicações P2P atravessaram problemas semelhantes, onde a inundação se tornou ineficiente na busca por informação. Estas aplicações evoluíram para um modelo baseado em DHT que possibilitou a distribuição de forma estruturada das informações da rede.

A presente proposta se baseou na arquitetura estruturada das aplicações P2P para propor um mecanismo eficiente de distribuição de informação da rede através



da adoção de um plano de informação. Este plano utilizou uma rede sobreposta para formar o plano de informação. Foi realizado neste plano o armazenamento estruturado das informações e a busca por estas. Uma reconfiguração determinística de topologia também foi apresentada para a rede do plano de informações para possibilitar a redução da ocupação dos enlaces por tráfego de controle e redução do atraso na atualização das informações.

A proposta atingiu o seu objetivo principal de apresentar uma nova solução à disseminação de informação em planos de controle de redes multi-camadas como observado nos avaliações apresentadas. Estas necessitam ainda de uma validação através de simulações computacionais porém, os resultados encontrados indicam perspectivas bastante animadoras para a utilização de técnicas P2P na disseminação de informação.

Dentre os estudos futuros necessários podemos descrever os seguintes itens:

- Para lidar com o atraso inserido enquanto se busca recursos no plano de informação, propõe-se inicialmente, uma análise mais elaborada da forma de construção do plano de informação, passando a considerar a disposição física dos nós. Para isso, o desenvolvimento de uma função de dispersão que agrupe nós próximos fisicamente deve ser considerado.
- A redução do número de nós no plano de informação também pode contribuir para a redução no tempo de busca, porém, estudos da confiabilidade e escalabilidade dessa alternativa devem ser feito pois há nesta proposta uma concentração das operações de controle do sistema tornando-o mais vulnerável à falhas e exigindo mais recursos de processamento e memória de cada nó.
- O estudo de um mecanismo híbrido que utilize inundação para divulgar infor-

mações que não sofrem alterações constantemente, como a topologia da rede e, utilize também, um sistema estruturado para disponibilizar informações mais dinâmicas, como capacidade de cada enlace e disponibilidade de comprimentos de onda. Esta estratégia permitiria reduzir a quantidade de buscas a serem realizadas no plano de informação sem a necessidade de atualizações constantes para manter a base de dados atualizada.

- Uma nova forma de distribuição de informação que ao invés de publicar o recurso como um todo publique as restrições atendidas pelos recursos, possibilitando assim buscar de uma só vez os recursos necessários ao estabelecimento do enlace fim-a-fim reduzindo o tempo de operação do sistema.
- Estudos sobre a integração da arquitetura de distribuição de informação com os protocolos especificados pelo IETF no GMPLS, principalmente os de reserva de recurso. Uma vez que existe um atraso maior no estabelecimento do LP, o que pode causar a reserva de muitos recursos por um longo tempo, o que pode aumentar a probabilidade de bloqueio do sistema.

## Referências Bibliográficas

- [1] J. Postel. Transmission control protocol - darpa internet program protocol specification,, 1981.
- [2] J. Postel. Internet protocol - darpa internet program protocol specification, rfc 791, 1981.
- [3] Sung Hwan Hwang, Jung Woon Lim, and Byung Sup Rho. 120 gb/s-level vcsel array optical subassembly using passive alignment technique, 27–30 May 2008.
- [4] F.E. Doany, C.L. Schow, C.K. Tsang, N. Ruiz, R. Horton, D.M. Kuchta, C.S. Patel, J.U. Knickerbocker, and J.A. Kash. 300-gb/s 24-channel bidirectional si carrier transceiver optochip for board-level interconnects, 27–30 May 2008.
- [5] K.M. Sivalingam and S. Subramaniam. Emerging optical network technologies: Architectures, protocols and performance, 2004.
- [6] D. Colle S. De Maesschalck, M. Pickavet and P. Demeester. Multi-layer traffic grooming in networks with an ip/mps layer on top of a meshed optical layer, 2003.
- [7] J. Awduche D. Rajagopalan, B. Luciani. Ip over optical networks: Architecture aspects, 2000.

- [8] J. Moy. Ospf version 2, 1998.
- [9] K. Yeung D. Katz, D. Kompella. Traffic engineering (te) extensions to ospf version 2, 2003.
- [10] D. Kamat S. Guerin R. Orda A. Przygienda T. Apostolopoulos, G. Williams. Qos routing mechanisms and ospf extensions, 1999.
- [11] A.V. Aho and D. Lee. Hierarchical networks and the lsa n-squared problem in ospf routing, 27 Nov.-1 Dec. 2000.
- [12] C.-X. Chi, D.-W. Huang, D. Lee, and X.-R. Sun. Lazy flooding: a new technique for signaling in all optical network, 17-22 Mar 2002.
- [13] G. Apostolopoulos. Forwarding plane flooding, 2003.
- [14] C. Pinart, R. Munoz, and G. Junyent. Experimental implementation of distributed management for service provisioning in an ason/gmpls testbed, 2004.
- [15] K. Khavari, N. Abji, R. Farha, Chuen Liang, A. Tizghadam, F. Fadaie, and A. Leon-Garcia. Structured peer-to-peer control plane, June 2006.
- [16] I. Stoica, R. Morris, D. Liben-Nowell, D.R. Karger, M.F. Kaashoek, F. Dabek, and H. Balakrishnan. Chord: a scalable peer-to-peer lookup protocol for internet applications, Feb. 2003.
- [17] Francis P. Handley M. Karp R. Ratnasamy, S. and S. Shenker. A scalable content-addressable network, 2001.
- [18] R. Ramaswami and K. Sivarajan. Optical networks: A practical perspective, 2001.

- 
- [19] J. Strand, A.L. Chiu, and R. Tkach. Issues for routing in the optical layer, 2001.
- [20] Raymond A. Patterson Richard Baar. Grooming telecommunications network, 2001.
- [21] Huaxiong Yao, Yan Yang, and Xiansi Tan. Resource efficient grooming policy in groomable wdm mesh networks, 28–30 April 2008.
- [22] Tibor Cinkler, Szilard Zsigmond, and Marcell Perenyi. Traffic grooming and power level tuning for physical impairment constrained routing, 22–26 June 2008.
- [23] O. Awwad, A. Al-Fuqaha, and A. Rayes. Performance of wdm mesh networks with limited traffic grooming resources, 2–4 July 2007.
- [24] A. Jaekel, A. Bari, Ying Chen, and S. Bandyopadhyay. New techniques for efficient traffic grooming in wdm mesh networks, 13–16 Aug. 2007.
- [25] F. Solano, L.F. Caro, J.C. de Oliveira, R. Fabregat, and J.L. Marzo. G+: Enhanced traffic grooming in wdm mesh networks using lighttours, June 2007.
- [26] Luis Gustavo Zuliani. Arquitetura e implementação de um serviço de informações topológicas e de engenharia de tráfego para sistemas rwa, 2006.
- [27] T. E. Stern and K. Bala. Multiwavelength optical networks: A layered approach., 1999.
- [28] N. Ghani, S. Dixit, and Ti-Shiang Wang. On ip-wdm integration: a retrospective, 2003.

- 
- [29] L. Raptis, G. Chatziliass, A. Manzalini, C. Cavazzoni, G. Ricucci, G. Lehr, U. Hartmer, R. Geerdsen, M. Jaeger, A. Gladisch, D. Colle, S. De Maesschalck, I. Lievens, P. Demeester, and L. Piergiovanni. Design and experiments of an automatic switched optical network (ason), 30 Sept.–4 Oct. 2001.
- [30] E. Mannie. Generalized multi-protocol label switching (gmpls) architecture, 2004.
- [31] J. Lang. Link management protocol (lmp), 2005.
- [32] A. Farrel and I. Bryskin. Gmpls: Architecture and applications, 2005.
- [33] Caixia Chi, Dawei Huang, D. Lee, and XiaoRong Sun. Lazy flooding: A new technique for information dissemination in distributed network systems, Feb. 2007.
- [34] T. Okazaki, H. Kamada, K. Kinoshita, H. Tode, and K. Murakami. Efficient flooding control suitable for multiple requests, 23-27 Jan. 2006.
- [35] A. Shaikh, J. Rexford, and K.G. Shin. Evaluating the impact of stale link state on quality-of-service routing, April 2001.
- [36] Hai Liu, Xiaohua Jia, Peng-Jun Wan, Xinxin Liu, and Frances F. Yao. A distributed and efficient flooding scheme using 1-hop information in mobile ad hoc networks, May 2007.
- [37] C. Chi, X. Sun, and Y. Qian. Evaluating the impact of flooding schemes on best-effort traffic, 28 Nov.-2 Dec. 2005.
- [38] L. Berger. Generalized multi-protocol label switching (gmpls) signaling resource reservation protocol-traffic engineering (rsvp-te) extensions, rfc 3473, 2003.

- 
- [39] Byrav Shen, Lu; Ramamurthy. Provisioning and restoration in the nextgeneration optical core, 2003.
- [40] R.A. Guerin and A. Orda. Qos routing in networks with inaccurate information: theory and algorithms, June 1999.
- [41] Y. Ohara, M. Bhatia, N. Osamu, and J. Murai. Route flapping effects on ospf, 27-31 Jan. 2003.
- [42] OpenNap. Napster protocol specification [online]. <http://opennap.sourceforge.net/>, 2001.
- [43] Gnutella. The gnutella protocol specification [online]. <http://www.clip2.com/Gnutellaprotocol04.pdf>, 2000.
- [44] Inc BitTorrent. Bittorrent [online]. <http://www.bittorrent.com/>, 2008.
- [45] Emule. The emule protocol specification [online]. <http://www.emule.com/>, 2005.
- [46] Team RetroShare. Retroshare: secure communications with friends. <http://retroshare.sourceforge.net/>, 2007.
- [47] Cspace. Secure, decentralized, user-to-user communication [online]. <http://cspace.in/>, 2008.
- [48] Rajaraman R. Plaxton, C. and A. Richa. Accessing nearby copies of replicated objects in a distributed environment, 1997.
- [49] Kubiatawicz J. Zhao, B. Y. and A. Joseph. Tapestry: An infrastructure for fault-tolerant wide-area location and routing, 2001.

- [50] Antony Rowstron and Peter Druschel. Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems, 2001.
- [51] M. Xie. P2p systems based on distributed hash table, 2003.
- [52] K. Gummadi. The impact of dht routing geometry on resilience and proximity, 2003.
- [53] T. Heer, S. Gotz, S. Rieche, and K. Wehrle. Adapting distributed hash tables for mobile ad hoc networks, 13-17 March 2006.
- [54] Venâncio Raposo Nelson Ciriz. Resolução de colisões. [on line] <http://w3.ualg.pt/hshah/ped/Aula2008>.
- [55] National Institute Standards Technology. Secure hash standard, August 2002.
- [56] R. Rivest. The md5 message-digest algorithm, 1992.
- [57] Mei Li, Wang-Chien Lee, and A. Sivasubramaniam. Semantic small world: an overlay network for peer-to-peer search, 2004.
- [58] Hui Zhang, A. Goel, and R. Govindan. Using the small-world model to improve freenet performance, 23-27 June 2002.
- [59] Stanley Milgram. The small world problem, 1967.
- [60] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks, 1998.
- [61] R. Axtell. Behavioral dimensions of retirement economics, 1999.
- [62] Francesc Comellas and Michael Sampels. Deterministic small-world networks, 2001.



- [63] Jon Kleinberg. The small-world phenomenon: An algorithmic perspective, 2000.
- [64] Joseph G. Peters Francesc Comellas, Javier Ozón. Deterministic small-world communication networks, 2000.
- [65] Marconi P. Fardin, Moisés R. N. Ribeiro, and Luis C. Calmon. Reconfiguração determinística de topologia para geração de característica “small world”, 2005.
- [66] Helio Waldman Marconi Pereira Fardin, Moisés R. N. Ribeiro. Small-world reconfigured can for distributed control plane of future optical networks, 2008.
- [67] S. McCanne and S. Floyd. The network simulator - ns-2. [on line] <http://www.isi.edu/nsnam/ns/>, October 2008.
- [68] T. Korkmaz and M. Krunz. Hybrid flooding and tree-based broadcasting for reliable and efficient link-state dissemination, 17–21 Nov. 2002.